Fourth Quarter of 2023
Issue No. 4
L023010001

R&A electronics

IC The World

Chip Revolution on AI Servers

www.randa.sg

World-Class Sourcing and Quality

R&A Blue Book Market Report

Industry Trends

Industry Focus

R&A electronics Marketing Insight

The update on the chip market trend in August

Highlights

- Industrial and automotive markets remain in growth
- The analysis of the R&A trading report

We look forward to providing you with profound insights and in-depth analysis of the semiconductor industry, helping you gain a better understanding of this ever-evolving field and discover the allure of cutting-edge technology. we are committed to offering you comprehensive information and support.

Collect and Analyze the Official Reports

IC The World is a showcase for the global economy and semiconductor industry. We analyze all relevant data to dissect the global economy and chip market. This issue of the journal aims at the trending sector--Artificial intelligence, analyzing the applications of AI in various areas and its influence on the whole market.

In-depth Research

We delve into diverse aspects of the semiconductor market, including technological trends, market share, and competitive landscape.

The Semiconductor Industry Association (SIA) released the total global semiconductor sales in the second quarter of 2023 of $124.5 billion, a sequential increase of 4.7% and edging up for four consecutive months.

Since cyclical recovery occurs, a resurgence in the semiconductor industry is emerging from the downturn.

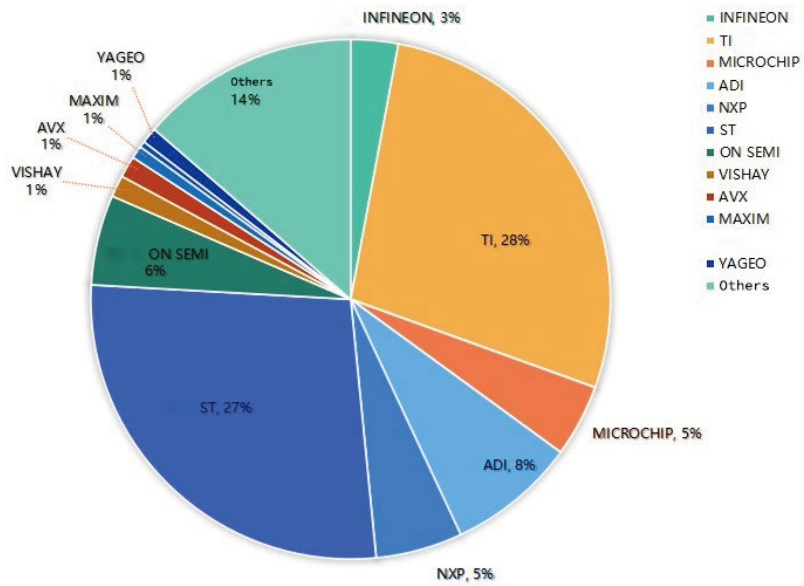
The areas of Industry and automobiles remain a growth highlight which are in great demand.
The result is represented by Various Q2 financial reports from chip giants.

Chipmakers	Q2Revenue ($0.1 billion)	Q1Revenue ($0.1 billion)	Q2 YoY	Q2 MoM	GROWTH	DECLINE
INTEL	129	117	-15%	10.20%	IFS foundry business stood out as the brightest highlight.	The four major sectors of Intel -- client computing, data center & artificial intelligence, NEX, and Mobileye -- all continued their year-on-year declining trend in revenue.
Samsung	111.31	103.75	-48%	7.30%	★ Demand for artificial intelligence applications maintains robust, resulting in an improved performance for the memory business compared to the previous quarter. ★ There is a primary focus on HBM and DDR5 products.	The mobile business saw substantial losses.
NVIDIA	135	71.92	101%	88%	Data centers, Gaming, and automobiles accomplish a leap thanks to the boom of demand for AI-related chips.	/
Broadcom	87.33	89.15	8%	equivalence	The growth rate for AI-related chip revenue is expected to highly promote.	/
Qualcomm	84.51	94.63	-22.70%	-8.90%	The automotive business is ascending to $434 million, with a 13% year-on-year rise.	Mobile and IoT demands remained sluggish.
SK Hynix	55.5	38.4	-47%	44%	★ Generative AI centered around ChatGPT is making its way into the whole technology, which is anticipated to push a surge in storage chip demand for AI servers. ★ Sales of premium products such as HBM3 and DDR5 are about to expand.	The storage market remained dull, with PC and smartphone markets stagnating, causing continued price drops for general DRAM products like DDR4.
AMD	53.6	53.53	-18%	equivalence	The embedded department benefited from strong growth in industrial, vision, healthcare, automotive, and testing & simulation markets.	AMD's three major departments - data center, customer business, and gaming - have experienced a decline.
TI	45.3	43.79	-13%	equivalence	The personal electronics and automotive chip markets have achieved growth.	Demand in other end markets, such as communication equipment, enterprise systems, and analog devices, has been lackluster.
Infineon	44.6	43.4	-0.70%	equivalence	The Automotive and Green Industrial Power (GIP) markets have significantly gained.	Power and Sensor Systems (PSS) witnessed a notable decrease.
ST	43.3	42.5	12.70%	equivalence	Revenue is constantly gaining reactions from the growth in automotive and industrial businesses.	Personal electronics business revenue has slashed.

Qualcomm reported Q2 2023 revenues of $8.451 billion, a 22.7% decrease from the $10.936 billion in the same period the previous year. The net profit was $1.803 billion, down by 51.7% year-on-year. Despite this poor performance, its automotive business remains the only segment in its core revenue that has sustained growth.
STMicroelectronics saw a 12.7% year-on-year increase in net income in the second quarter, beneficial from the robust growth of its automotive and industrial businesses offsetting the decline in electronics.
AMD's revenue for Q2 2023 remained consistent with the previous quarter but plummeted by 18% year-on-year. On one hand, there was a drop in its data center, customer business, and gaming departments. On the other hand, the embedded department benefited from strong growth in industrial, vision, healthcare, automotive, and testing & simulation markets, with revenues soaring to $1.5 billion, a 16% spike year-on-year.

Based on delivery records of R&A Electronics in July, the performance of Texas Instruments' analog devices and STMicroelectronics' discrete devices has been notably impressive.

Here's an overview of the semiconductor market trends.



The sectors of industrial control, automotive, photovoltaic, and energy storage have been barely stagnant, but moving ahead steadily.

Automotive MCUs of NXP's MCFXX series, ST's STM32 series, and TI's TMS320 series enjoy a rising popularity.

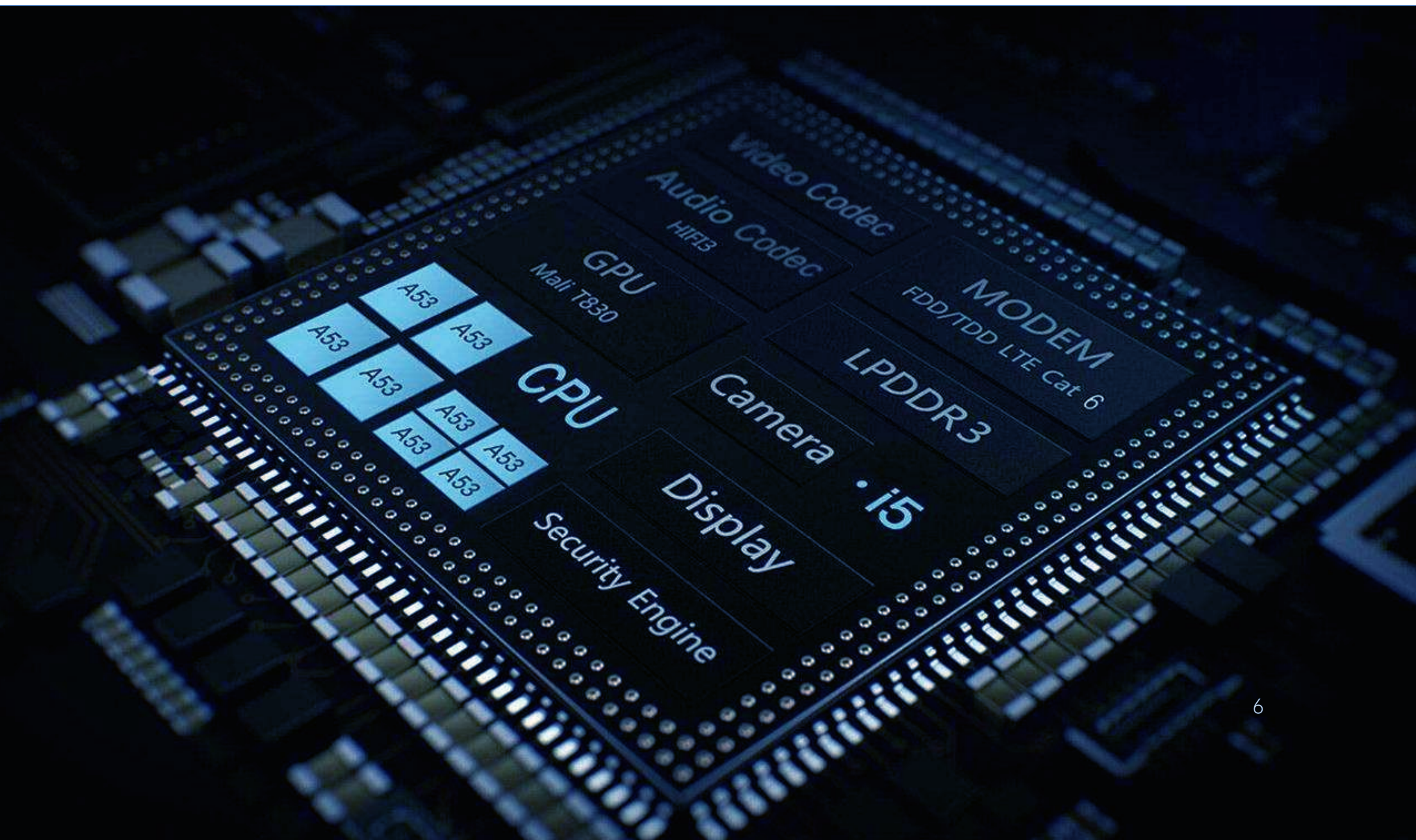
	Manufacturer	catogary	Industry
	TI	32位 MCU 32-bit MCU	Consumer
	NXP	32位 MCU 32-bit MCU	Automotive
		32位 MCU 32-bit MCU	Industrial
MCU	ST	32位 MCU 32-bit MCU	Consumer
		8位MCU 8-bit MCU	Battery
	MICROCHIP	8位MCU 8-bit MCU	Industrial
		8位MCU 8-bit MCU	Others

The lead time for analog devices as a whole is stable yet declining, especially general-purpose analog devices, while automotive analog and switching regulators still have lead times of over 40 weeks.

	Manufacturer	catogary	Industry
Analog Devices	TI	Power Management（Voltage Reference)	Industrial
	TI	Temperature Sensor	Industrial
	TI	Operational Amplifier	Industrial/consumer
	ANALOG DEVICES	Switching Regulator	Industrial
	TI	Precision Amplifier	Industrial
	NXP	CAN Interface Integrated Circuits	Automotive
	TDK	NTC Thermistor	Consumer
	ON SEMI	Operational Amplifier	Automotive
	Others	NTC Thermistor	Consumer
	INFINEON	Controller/Driver	Automotive
	ADI	ASRC	Consumer
	ST	Gate Driver（PMIC)	Automotive
	Others	Thin Film Resistor	Industrial
	TI	Current-sensitive Amplifier	Automotive
	TI	Analog to Digital Converter	Industrial/Automotive
	TI	Switch Controller	Automotive
	Others	Switching Regulator	Consumer
	MAXIM	Operational Amplifier	Industrial
	TI	Switches and Multiplexers	Automotive
	TI	Supervisor	Industrial
	TI	Power Management（Voltage Reference)	Others
	TI	Switching Regulator	Industrial
	NXP	Audio amplifier	Automotive
	TI	Switching Regulator	Consumer
	TI	Switching Regulator	Automotive
	INFINEON	Touch Screen Controller	Automotive
	PARADE	Interface IC	Telecom
	MAXIM	Switch IC	Consumer
	ST	Gate Driver（PMIC)	Automotive
	Others	Switching Regulator	Battery
	NXP	USB Interface Integrated Circuits	Consumer
	INFINEON	Power Switch	Automotive
	Others	Sensor	Consumer
	ST	Gate Driver（PMIC)	Automotive
	Others	Pressure Sensitive Resistor	Industrial
	TI	Controller/Driver	Industrial
	TI	Converter	Consumer
	ANALOG DEVICES	Digital PotentiometerIndustrial Resistance	Others
	ANALOG DEVICES	PMIC	Others
	TI	Temperature Sensor	Others
	Others	Switching Regulator	Automotive
	Others	IC Switch IC	Industrial
	TI	Interface IC	Industrial
	NXP	SoC	Consumer
	Others	USB Switch（PMIC)	Automotive
	MICROCHIP	Touch Screen Controller	Consumer

In the second quarter of this year, global DRAM sales experienced an increase, signaling a resurgence in the DRAM market.
High-end semiconductors such as HBM have become the next strategic focus for major memory chip manufacturers.

	Manufacturer	catogary	Industry
Memory	SAMSUNG	DRAM	Consumer
	MICROCHIP	EEPROM	Industrial
	KIOXIA	NAND Flash Memory	Consumer
	INFINEON	NVRAM	Industrial
	GIGADEVICE	NAND Flash Memory	



Both ST and Infineon report a general lead time ranging from 40 to 52 weeks for their discrete devices without further improvement. This includes devices like high and low-voltage MOSFETs, IGBTs, wide bandgap MOSFETs, rectifiers, and bipolar transistors.

	Manufacturer	catogary	Industry
OTHERS	ADI	LED Lighting Driver	Consumer
	TI	Voltage Converter	Industrial
	TI	LED Lighting Driver	Automotive /Consumer
	MICROCHIP	FPGA	Medical
	NEXPERIA	Logic Gate	Industrial

	Manufacturer	catogary	Industry
Discrete Devices	ST/ONSEMI	Schottky Diode Rectifier	Industrial/Consumer
	INFINEON	MOSFET	Industrial/Automotive
	MICROCHIP		
	ST		
	VISHAY		
	LITTELFUSE		
	IXYS		
	TI	Low-Dropout Regulator	Industrial/Automotive
	ONSEMI		
	ST		
	VISHAY		
	BOURNS		
	COILCRAFT		
	INFINEON		
	VISHAY	TVS Diode	Consumer
	ST	Bipolar Transistor	

There has been a surge in demand and prices for PCIe Switches, Ethernet chips, and HBMs, which are used in AI servers. For instance, Broadcom's SS series of PCIe Switch chips saw an increase of over 200% in price. Additionally, its Ethernet chips for automotive and AI applications witnessed a spike of over 100% in price.

	Manufacturer	catogary	Industry
RF/Wireless	TAIYO YUDEN	Duplexer	
	SKYWORKS	RF Module	Consumer
	NXP	LIN Transceiver	Automotive
	BROADCOM	Ethernet IC	Consumer
	INFINEON	RF Receiver	Consumer

The move towards intelligent automobiles and the AI server trend has brought a jolt of new demand for passive components, spurring a recovery in passive device sectors.

	Manufacturer	catogary	Industry
	YAGEO	Thick Film Resistors- SMD	Automotive /Consumer/Others
	KOA		
Passive devices	VISHAY		
	TDK	MLCC	Automotive /Consumer/Others
	KEMET		
	AVX		
	KEMET	Tantalum capacitor - Solid SMD	Consumer
	COILCRAFT	Power inductor	
	BOURNS		

In summary, even though the global economy remains depressed and the consumer electronics market is stalked by uncertainties, the robust demand in the automotive and industrial fields, combined with the booming AI trend in 2023, has lent an impetus to the somewhat languishing semiconductor industry. In the revenue outlook for the third quarter, all major chip-makers maintain a positive stance, indicating that demand for semiconductors may gradually pick up.

R&A electronics

Anti-Counterfeiting Testing Laboratory

Analysis of the Quality Inspection Report and the Chip Market for August

Currently, the chip trading market is plagued by various deceptive practices, from counterfeit labels to substandard chips, causing significant losses for many individuals and businesses.

Through counterfeit cases detected by R&A Laboratory, we aim to demonstrate how we ensure product quality through a multi-layered protection process to safeguard our customers' products. Additionally, we want to raise awareness about which products are more susceptible to counterfeiting.

Highlights

- Sharing Counterfeit Cases
- Market Analysis | Trends and Changes in the chip market

First-tier interception

The packaging label of the product is checked by comparing it with the label and packaging of the manufacturer that are recorded in our label database. This step can verify and prohibit 90% of counterfeit products.

Second-tier interception

Based on IDEA-STD-1010B testing standards, the inspectors observe the surface, markings, and leads of chips by digital microscope and proceed with dimensions measurement, acetone, and scratch test to verify if there is oxidation, scratches, resurfacing, plating or contamination on the chips.

Third-tier interception

Deep inspections are executed to observe the internal structure and elements of chips, involving X-RAY, decapsulation, solderability, cross-section, and electrical test

Cases of Fake products

Apart from presenting cases of counterfeit products encountered in the lab, the article offers chip market analysis and trending areas for August based on our received goods. In August, products from 55 manufacturers were received in the R&A warehouse, in which the most prominent quantities were TI, Onsemi, NXP, Microchip, Infineon, ST, and ADI.

Here are examples showcasing counterfeit labels and visual details that did not pass the inspection. Please take a look at the cases.

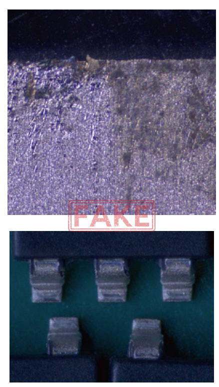
First-level interception

Manufacturers	Part model	Issue
ONSEMI	MBRM140T1G MBRA160T3G 1SMB5924BT3G	Fraudulent labels. The positions of letters are little in line with the specifications of manufacturers.
TI	UCC28089DR	
MICROCHIP	PIC16F677T-E/SS	
NXP	TJA1040T/CM	

Here are examples showcasing counterfeit labels and visual details that did not pass the inspection. Please take a look at the cases.

Second-level interception

Manufacturers	Part model	Issue
ONSEMI	Part model	The metal part is acutely oxidizing.
ADI	FCH023N65S3-F155 ISL9R1560G2	
	ADP123AUJZ-R7	



In spite of Texas Instruments' recent disappointing financial reports and orders, its automotive business remains impressive and its industrial chips also consistently perform well. Based on Onsemi's Q2 financial report, its business has skyrocketed, in the face of revenues on a par with last year. Notably, the automotive sector has ballooned up 35% and their industrial business up 5% from a year earlier. However, the upcoming months will be challenging for the spot market since some factories will seek long-term order opportunities to reduce costs.

Onsemi's silicon carbide business has soared, almost quadrupling YoY, because the upsurge of vehicle electrification and intelligence, Industrial automation, and energy storage like photovoltaics, has directly spurred one that makes the firm secure a substantial market share. Therefore, Onsemi succeeded in laying a solid foundation for future endeavors in the silicon carbide domain. Supposedly, its orders in August are on the rise with a low inventory. The primary demand in August was for automotive and industrial products, such as the NCV and SZ series, the company said. Delivery times for automotive chips remain at 40-50 weeks with no significant improvements. Moreover, the popular FSV series has a delivery period of about 50 weeks, with persistently high market prices.

As for NXP Semiconductors, while partial industrial and automotive chips have seen improved delivery times, the demand still outstrips the supply, particularly for the highly sought-after MC52xx and S912ZVxx. The general-purpose MCU, Kinetis K series, remains scarce.

Since last year, demand for Ethernet switch ICs of Microchip has skyrocketed, with certain KSZx series still out of stock and priced high. Infineon saw a decrease in automotive demand in August, as the SAK series experienced a fierce price drop. The supply of regular high and low-voltage MOSFETs is increasingly adequate. By contrast, some high-voltage MOSFETs still have high prices. IGBTs remain out of stock, posing extended delivery times of over 40 weeks for the most part, and the costly spot market.

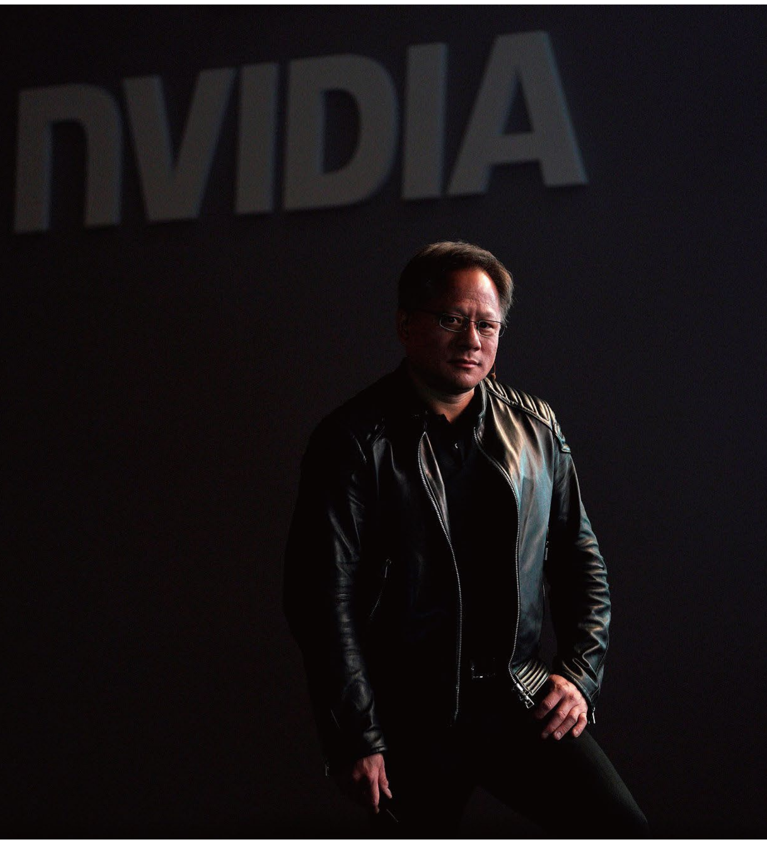
STMicroelectronics showed a stable order volume in August and a low inventory; automotive MCU lead times were kept between 40-52 weeks. Xilinx's XCF series PROM has been discontinued, but customer demand persists, which causes a slight increase in market prices. Generally, Xilinx indicates that delivery times are gradually improving as a whole, although, that of the 6S series stays unchanged.

Broadcom looks at the AI-related field. Its upscale PLX chips including SS26 and SS24 are in short supply, posing a headache to AI server shipments and the market pricing. Renesas' customer interests mainly centred around automotive products, like automotive MCUs, CLKs, memory, power management, etc. Separately, product codes starting with R5, R5f, and R7FUPD suffer the short supply and high market prices.

Apparently, the automotive industry is moving ahead, and the industrial sector making steady progress. As renewable energy is continuously pushed forward, China and the US maintain growth in photovoltaic storage. Given declining electricity prices in Europe, the growth rate of photovoltaic storage has slowed down, but top storage suppliers in Europe continue to see rising shipments.

Emerging markets in South Africa, Southeast Asia, and Latin America are all stepping up. From 2020-2022, South Africa's proportion of renewable energy generation expanded from 5.6% to 7.3%. Vietnam, stimulated by the FIT policy, has accomplished a surge in solar installations, with a total installed capacity above 19GW by the end of 2022, leading in Southeast Asia. Other countries in Southeast Asia have also introduced incentive policies to foster photovoltaic storage and other energy developments. Thailand has implemented a new FIT from 2022-2030, with battery storage at 2.83 THB/kWh. The Philippines now allows overseas vendors to own 100% of renewable energy assets. Malaysia aims for renewable energy to account for 70% by 2050, and Indonesia plans to rely entirely on renewable energy for power by 2060. Australia's energy storage is frenziedly ascending, owing to the energy crisis and policy subsidies.

AI servers see a giant hit with soaring prices and popularity.

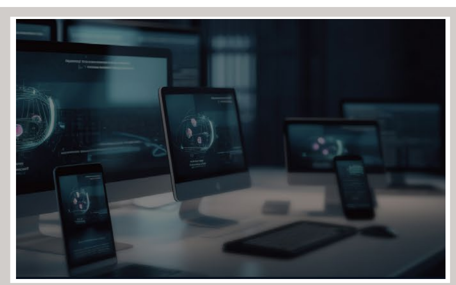


NVIDIA CEO Jensen Huang stated, "The worth-around-$1 trillion data center which has been installed across the world is undergoing a shift away from general computing to the accelerated and generative artificial intelligence. A new era of computing has opened." In this transformation, AI servers are indispensable.

what are AI servers?

So, what are AI servers? The biggest difference between regular servers and AI servers is that the latter ones typically employ "the combination" such as CPU+GPU, CPU+TPU, CPU+other acceleration cards, etc, of which the CPU completely sheds the burden of computation

Therefore, AI servers usually feature several GPUs and only 1-2 CPUs. Depending on the number of GPUs, AI servers can be further categorized as four-way, eight-way, and sixteen-way servers, with the eight-way AI server equipped with 8 GPUs being the most popular.



The cost breakdown of essential components in AI servers is as follows: GPUs account for 48%, CPUs for 7%, storage memory (HBM) for 9%, DRAM for 9%, SSD+RAID storage for 10%, interconnection interface chips for 0.7%, network cards for 1%, power supply components for 0.09%, BIOS/BMC chips and firmware for 0.02%, and power modules for 0.3%.

As high computing is popularly required, GPUs are the favored choice. Although FPGAs and various ASIC chips can perform extensive calculations, GPUs possess remarkable versatility. According to IDC data, in the first half of 2021, GPUs significantly became the preferred choice for accelerating data centers in the Chinese artificial intelligence chip industry, capturing over 90% of the market share.

Since the launch of ChatGPT by OpenAI earlier this year, which supports both image and text outputs and boasts powerful image recognition capabilities, it has set the entire tech industry abuzz. Subsequently, the large model market is flooded with tech titans like OpenAI, Google, Microsoft, Amazon and Chinese giants such as Baidu, Alibaba, Huawei, and Tencent. As of now, China and the United States win an 80% slice of large models worldwide, with the United States leading the world in the number of that.

AI服务器的产业链

Upstream	Manufacturers	Applications
GPU/CPU/AI accelerator cards INTEL, NVIDIA, AMD, Cambrian, Hygon Information Technology Co., Ltd. **DRAM** Samsung, SK Hynix, Micron, Montage Technology, GigaDevice, Shenzhen Kaifa Technology Co., Ltd., Wuxi Taiji Industry Co., Ltd. **HDD/SSD** WesternDigital、Seagate、Toshiba、Samsung、SK Hynix、Micron、Biwin、Shenzhen Longsys Electronics Co.,Ltd **Interconnection interface** NVIDIA, INTEL, TI, ADI, Rambus, Renesas, Parade, Montage Technology **Network Adapter** Mellanox **AI server Battery** Delta, Emerson, Liteon, GreatWall, Xinleineng, Beijing Dynamic Power Co., Ltd. **discrete/ passive devices** Joulwatt, Shanghai Bright Power Semiconductor, Poco Holding	**ODM** Foxconn Industrial Internet Co., Ltd. Inventec Wistron Quanta Wiwynn Mitac Gigabyte Pegatron **Sever manufacturers** Inspur Sugon H3C Lenovo HPE DELL	Internet/ Cloud service providers Google Aws Microsoft Facebook IBM Alibaba Baidu Tencent Mobile operators China Mobile China Telecom China Unicom China Broadcasting Network Corporation Enterprise market Government customers Financial / Healthcare / Energy, etc. Small and medium-sized enterprise (SME) customers

AI servers' popularity has quickly led to a shortage of AI-related chips such as GPUs. In terms of a report released by TrendForce, American cloud service providers (CSPs) began setting up Surface Mount Technology (SMT) production lines in Southeast Asia in late 2022, aiming to control the core supply chain.

Google, leading in the development and supply chain of its self-designed TPU machine, primarily holds the reins of both aspects. The production hub is also centered around Google's manufacturing facilities in Thailand. However, they still rely on Original Design Manufacturers (ODM) for workforce scheduling and production planning, while Google maintains control over other materials. To avoid supply chain disruptions, Microsoft, Meta, and AWS have not only focused on supply chain flexibility but also integrated system integration companies in the ODM production process for a more diversified and meticulous project execution.

Not only Apple, but companies like Amazon, AWS, Google, and others are also making significant advancements in the generative AI field

TrendForce indicates that Taiwanese server ODMs, involving Quanta, Foxconn, Wistron (including Wiwynn), and Inventec, have production facilities in Thailand, Vietnam, Malaysia, and other locations. It is estimated that the SMT production capacity, by 2023, in these regions will take about 23%, rallying to nearly 50% by 2026.

The move of CSPs to control the core of the supply chain has presented splendid opportunities for ODMs in the AI server market. Rumors suggest that Foxconn has exclusively secured a multitude of AI server orders from Apple, which are being shipped from their Vietnamese facilities. Apart from Apple, Amazon, AWS, Google, and others are making efforts in the generative AI field, pushing a substantial spike in orders for Foxconn's server business.

This has had a profound impact on AI server shipments, impeding downstream cloud service providers and AI companies. Prompted by the overwhelming demand for AI chips in data centers, NVIDIA intends to augment the production of A100, H100, and other GPU accelerator cards in 2024 to meet the enormous market demand.

NVIDIA inclines to triple the production capacity of GH100 chips, with shipments expected to reach between 1.5 to 2 million units next year, far surpassing this year's 500,000 units, insiders revealed.

Currently, NVIDIA's GPU chips are severely in short supply, triggering a surge in prices.

Memory chips, which occupied approximately 30% of the cost of AI servers, seem to be the first to rebound in the wave of artificial intelligence. According to Micron, AI servers require eight times the capacity of DRAM and three times of NAND compared to traditional servers, which is expected to drive a several-fold rise in server storage demand. For instance, Data center storage demand accounts for around 31% of the total storage market.

Memory chips, which occupied approximately 30% of the cost of AI servers, seem to be the first to rebound in the wave of artificial intelligence. According to Micron, AI servers require eight times the capacity of DRAM and three times of NAND compared to traditional servers, which is expected to drive a several-fold rise in server storage demand. For instance, Data center storage demand accounts for around 31% of the total storage market.

The global market scale of NAND Flash witnessed a sequential growth of 5% to $9.128 billion, while the DRAM market achieved a hike of 11.9% to $10.675 billion from the previous quarter. Generally, the worldwide memory-chip market accomplished $19.803 billion for the send quarter, up 9% month on month.

TrendForce claimed that the Q2 average fixed transaction price for 32GB DDR5 DRAM is estimated to ascend from $75 to $80-90. Analysts point out that 128GB DDR5 used in AI servers is ten times pricier than 64GB DDR4, and orders continue to mount.

Interface chip values are soaring rapidly as well thanks to the bump of AI servers. There are four types of interface chips used in AI servers: PCIe Switch, PCIe Retime, high-speed serial interface chips, and memory interface.

Firstly, PCIe Switch chips are essential for interconnecting PCIe devices in servers, including the interconnection between CPU and GPU. Scarce are the high-end PCIe 4.0 chips but with numerous demands. Broadcom's PCIe Switch chip, represented by the SS26, has seen an exceptional bloom in demand recently, with limited availability and prices surging from $5,500 to $20,000! Nowadays, mere a few companies worldwide possess manufacturing technology of PCIe Switch chips, such as Broadcom, Microchip, and ASMedia, which account for approximately 58% of the global market share.

Secondly, there are analog mixed-signal chips, Retimers. AI servers need at least a Retimer to ensure signal quality when connecting GPUs and CPUs. Most AI servers are equipped with multiple Retimer chips. For example, Astera Labs configures four Retimer chips in AI accelerators. At present, three major players are Parade, Astera Labs, and Montage Technology in the PCIe Retimer market.

Thirdly, high-speed serial interface chips contain NVIDIA's NVLink, AMD's Infinity Fabric, and Intel's CXL (Compute Express Link).

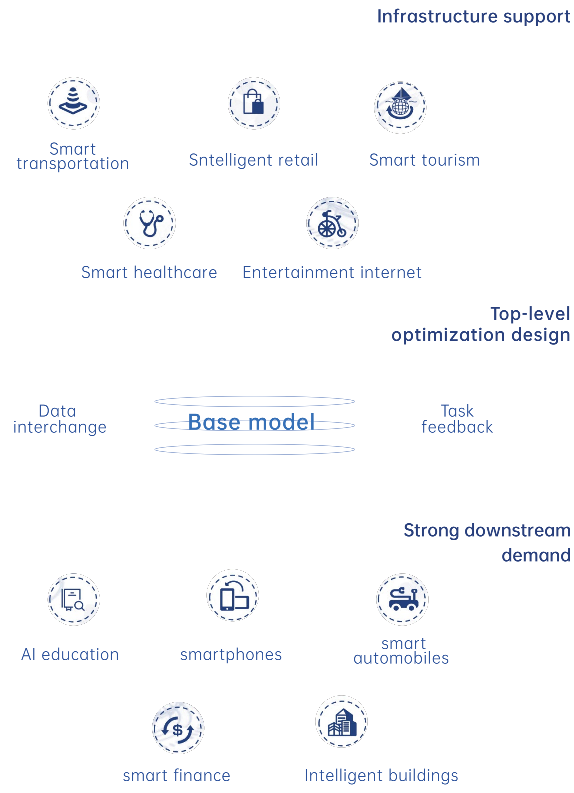
Lastly, memory interface chips cover Registering Clock Driver (RCD) and Data Buffer (DB). Memory interface chips are a technology-intensive industry that requires comprehensive and rigorous validation by manufacturers of CPU and memory, and OEM before they can be used on a large scale. It is demanding for new players to enter this market. In the DDR5 interface market, only Montage, Renesas Electronics (IDT), and Rambus provide mass-produced products for the first generation of DDR5 globally.

The value of chips related to AI server power modules is promoted three times. AI servers have power requirements that are 6-8 times higher than regular servers, which also increases the demand for power supplies. According to a research report released by Hu Yang, an analyst at Huaan Securities, on April 5th, regular servers used to be installed with only 2 units of 800W server power supplies, but AI servers now demand 4 units of 1800W high-power supplies. As a result, the cost of server power supplies has skyrocketed from $425 (3,100 yuan) to $1,700 (12,400 yuan), hiking threefold.

High-end optical module products are also in high demand in the era of AI. ZHONGJI INNOLIGHT CO., LTD., a leading optical module manufacturer, claimed that the second-quarter orders featured high-end products such as 800G/400G/200G. The company stated that some overseas customers began gradually entering the market of 800G in the fourth quarter of 2022, and thus an increasing number of data center customers are expected to start marching the field of 800G in 2024, stimulating the demand for 800G modules. InnoLight has fully mass-produced 800G optical modules, with Google and NVIDIA as major customers. The orders in the second quarter stood at 800,000 units. Alongside this, Meta may place orders in the third quarter of this year. Since March, there were additional orders for 100,000 units of 800G optical modules, followed by a 130,000-unit order added in April. By the end of May, Google ordered 150,000 units of optical modules again, totaling 600,000 units. Also, NVIDIA placed an extra order for 800G optical modules to 500,000 units. Therefore, it is anticipated that the total shipments of 800G optical modules in 2023 will hit between 1.2 million and 1.4 million units.

AI Empowering Automobiles: Spurring Automotive Intelligence!

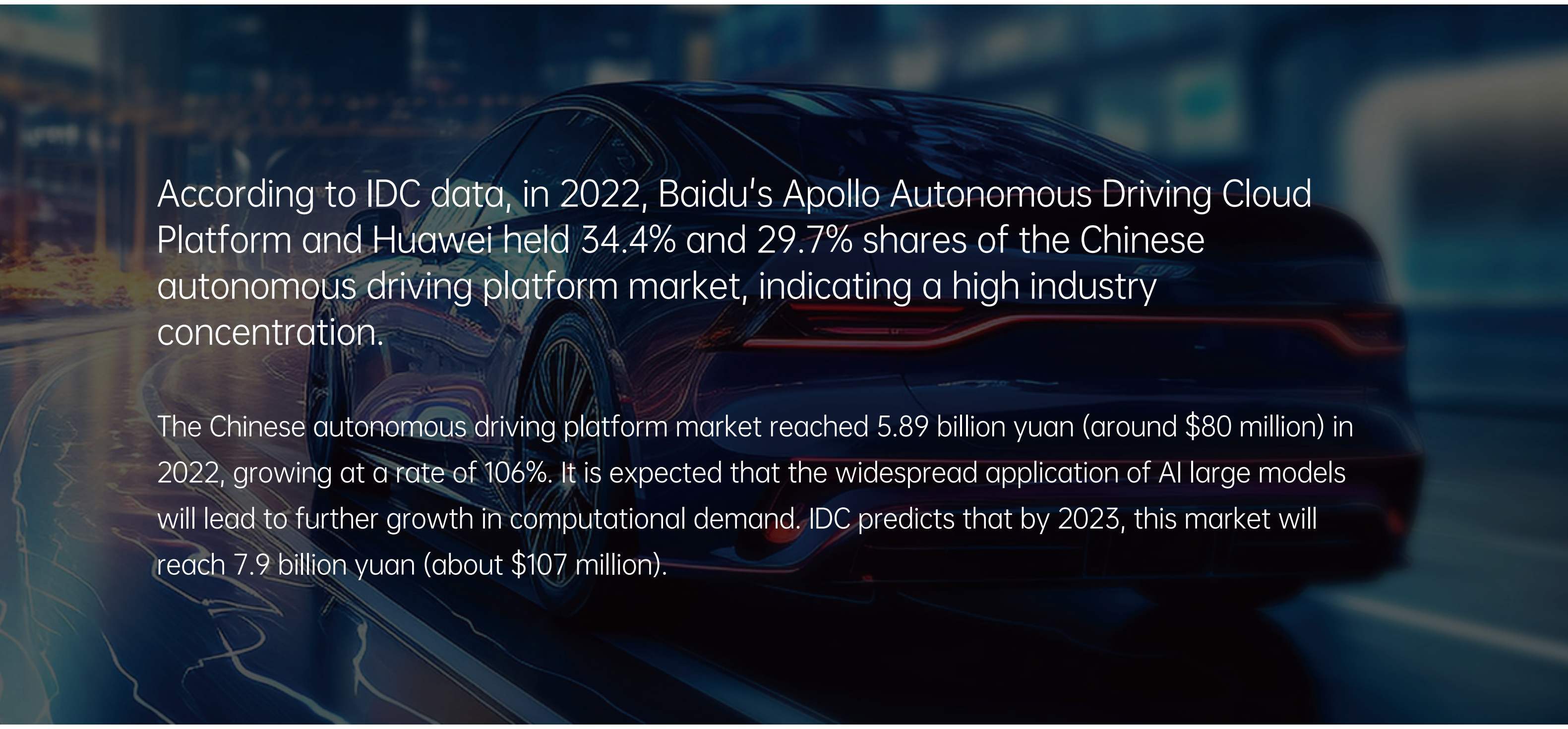
In November 2022, the American startup OpenAI rolled out ChatGPT, a chatbot powered by large-scale AI models. After its release, the platform experienced an explosive surge in user engagement, with over a million users within the first five days and exceeding a hundred million users within two months. This unprecedented user growth was fueled by ChatGPT's demonstrated ability in continuous dialogue, contextual comprehension, inquisitive reasoning, and the capacity to discern user intentions. It showcased the remarkable potential of artificial intelligence, prompting global companies to join the AI large model race. The scale of model parameters and computational demands is undergoing an exponential increase, as there is an active exploration of commercial applications for large AI models.



As foundational technologies undergo progressive innovation, base models and domain-specific large models continue to improve. The AI large model applications are poised to spread into various industries, including transportation, healthcare, and finance. It is in the cards to unleash a new wave of intelligent revolution led by superior and general artificial intelligence, which is about to significantly enhance production and efficiency. The economy, society, and industries will see an impressive reformation.

For instance, within the realm of comprehensive transportation, artificial intelligence has the potential to create value in various aspects of the automotive industry, like autonomous driving, intelligent cockpits, and operational management optimization. In the field of healthcare, artificial intelligence has the capacity to improve the efficiency of disease detection and facilitate intelligent analysis of medical imaging.

AI + Autonomous Driving: Intelligent algorithms shaping mobility scenarios.



According to IDC data, in 2022, Baidu's Apollo Autonomous Driving Cloud Platform and Huawei held 34.4% and 29.7% shares of the Chinese autonomous driving platform market, indicating a high industry concentration.

The Chinese autonomous driving platform market reached 5.89 billion yuan (around $80 million) in 2022, growing at a rate of 106%. It is expected that the widespread application of AI large models will lead to further growth in computational demand. IDC predicts that by 2023, this market will reach 7.9 billion yuan (about $107 million).

Human driving is based on cognitive understanding, relying on visual perception and the brain to make a decision. In contrast, autonomous driving struggles to match human-level driving perception, prediction, and decision-making capabilities, especially in complex and dynamic environments. AI large models are the product of "big data + high computational power + robust algorithms" and encapsulate the inherent wisdom of large datasets.

Therefore, AI large models are able to, through training, capture patterns in extensive data, leverage the advantages of increased model parameters, compensate for the accuracy shortcomings in autonomous driving, reduce in-vehicle inference computation time, and enhance automotive safety.

The construction of AI large models for autonomous driving is primarily divided into three modes: self-establishment of intelligent computing centers, independent training of large models as exemplified by Tesla; collaborative construction, with internet technology companies like Baidu and Tencent offering to-business intelligent automotive cloud services for OEMs to use in model training; solution partnerships. Large models like Baidu' s ERNIE Bot and SenseNova can be integrated into the vehicles to optimize the models for automotive applications.

Intelligent Cockpit + AI: New Form of Intelligent Space

According to IHSMarkit, the global market size for intelligent cockpits is projected to grow from $42 billion in 2021 to $68.1 billion in 2030, with a CAGR of 5.52% from 2022 to 2030. IHS predicts that the domestic market for intelligent cockpits in China will outpace global growth, with the domestic market size increasing from $9.9 billion in 2021 to an estimated $24.7 billion in 2030, with a CAGR of 10.69% from 2022 to 2030. China's share of the global intelligent cockpit market is expected to rise from approximately 24% in 2021 to 36% in 2030.

Furthermore, automotive intelligence is also reflected in the continuous enhancement of cockpits. With the accelerated integration of AI large models, intelligent cockpits have a significantly larger scope for evolution. They will no longer rely on improving the penetration rate of basic functions but will create a differentiated experience in a new form of intelligent space.

Currently, intelligent cockpits have evolved from traditional mechanical dashboards and car radios to intelligent assistant cockpits that feature biometric recognition and driver health monitoring. They are transforming into a third living space that integrates rich information, entertainment, and multifunctional capabilities. In the future, AI voice assistants will transition into "AI intelligent butlers," and automotive cockpits will actively provide drivers and passengers with more contextual services, enabling deep human-machine emotional interaction.

In March 2023, Baidu published the ERNIE Bot NLP large model, with its intelligent voice interaction functionality integrated into the Apollo intelligent cockpit system. In May 2023, iFLYTEK released the Xinghuo cognitive large model, announcing "large models + intelligent cockpits" as one of its four major industry applications. SenseTime also proposed products related to the automotive industry based on its large model system.

Nowadays, the penetration of intelligent cockpits has shifted from the initial dominance of "new energy vehicle" manufacturers to more collective progress across the entire industry. Several automotive companies, including Changan, JIDU, Geely, Voyah, Hongqi Auto, Great Wall, Dongfeng Nissan, and Leapmotor, have announced their integration with ERNIE Bot. Changan's Yidah will become the first model to incorporate ERNIE Bot and will be deployed in new vehicles through software upgrades. In May, Huawei stated that the AITO Quest M9 will be equipped with an AI large model, enhancing the in-car AI experience of Xiaoyi's intelligent assistant. iFLYTEK's "Xinghuo Cognition" large model also offers products related to the automotive sector, enabling cross-business and cross-scenario human-vehicle communication in the car.

The construction ofIntelligent computing centers of carmakers and cloud service providers

company	Intelligent computing center	Release time	Computing (PFLOPS)	Function
Tesla	Dojo intelligent center	August 2021	1800	Self-driving model training
Li Auto	Shanxi intelligent center built jointly with Volcengine	June 2023	At least 47 PFLOPS	
NIO	Partnering with Alibaba to make a "NIO Cloud" center			Autonomous driving, battery cloud service
Xpeng	"Whirlwind" computing center based on Alibaba cloud	August 2022	600	Self-driving model training
HAOMO.AI	MANA OASIS	January 2023	670	Self-driving model training
Geely	Geely Wise Star-Dubhe	January 2023	810	Internet of vehicles, self-driving, etc
SenseTime	SenseTime AIDC	January 2022	3740+1170 (in plan)	self-driving
Baidu	Kunlunxin	September 2022	200	self-driving
	Yangquan intelligent computing center	December 2022	4000	AI, Autonomous driving

Globally, the United States leads in the number of large models.

As early as 2020, the United States launched 15 large models, including GPT-3. China, starting in 2020, entered a period of rapid development for large models, introducing 30 large models in 2021, 28 in 2022, and 19 in the first five months of 2023, surpassing the United States' 18 large models. Currently, the quantity of large models in both China and the United States accounts for nearly 90% of the global total, with China's large model quantity ranking among the top.

In this fiercely competitive race, large models both domestically and internationally are advancing towards more refined technological directions, driving the overall technological advancement and practical applications of the AI industry, delivering faster and better empowerment to the automotive and other industries.

AI Boom pushes Medical Intelligence

In the era of digitalization, the rapid advancement of AI technology has brought about profound transformations across various industries. The healthcare sector, intimately tied to human life and well-being, stands at the forefront of this technological revolution.

The Intersection of AI and Healthcare

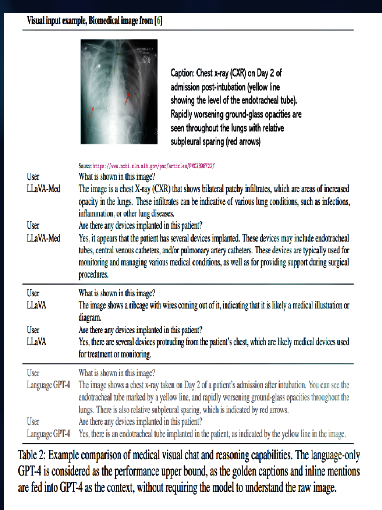
Since the inception of computers, scientists have been exploring their applications in the field of healthcare. With the enhancement of computing power and the reduction in data storage costs, medical data has become extensively digitized, laying the groundwork for the utilization of AI technology.

Technological Drive: In recent years, with the development of big data, deep learning, and cloud computing technologies, medical data has evolved from being a passive storage entity to actively training AI models. This enables early disease diagnosis and optimization of treatment plans, among other applications.

Market Demand:The global population is aging, and the prevalence of chronic diseases is increasing, posing greater demands on healthcare resources. AI technology's integration can help healthcare institutions use resources more efficiently to meet market needs.

The AI large models are evolving from general-purpose AI into specialized models for the healthcare industry

Microsoft has recently introduced the LLaVA-Med model, which is the first example of an effective multi-modal conversational AI model for the biomedical field. This model can answer open research questions related to biomedical images, providing robust support for AI applications in the medical field. In addition, domestic company United Imaging Intelligence has also explored training modes for medical AI large models, offering possibilities for the realization of interactive medical AI applications.


Visual input example, Biomedical image from [6]

Caption: Chest x-ray (CXR) on Day 2 of admission post-intubation (yellow line showing the level of the endotracheal tube). Rapidly worsening ground-glass opacities are seen throughout the lungs with relative subpleural sparing (red arrows)

Source: https://www.ncbi.nlm.nih.gov/pmc/articles/PMC7398722/

User	What is shown in this image?
LLaVA-Med	The image is a chest X-ray (CXR) that shows bilateral patchy infiltrates, which are areas of increased opacity in the lungs. These infiltrates can be indicative of various lung conditions, such as infections, inflammation, or other lung diseases.
User	Are there any devices implanted in this patient?
LLaVA-Med	Yes, it appears that the patient has several devices implanted. These devices may include endotracheal tubes, central venous catheters, and/or pulmonary artery catheters. These devices are typically used for monitoring and managing various medical conditions, as well as for providing support during surgical procedures.
User	What is shown in this image?
LLaVA	The image shows a ribcage with wires coming out of it, indicating that it is likely a medical illustration or diagram.
User	Are there any devices implanted in this patient?
LLaVA	Yes, there are several devices protruding from the patient's chest, which are likely medical devices used for treatment or monitoring.
User	What is shown in this image?
Language GPT-4	The image shows a chest x-ray taken on Day 2 of a patient's admission after intubation. You can see the endotracheal tube marked by a yellow line, and rapidly worsening ground-glass opacities throughout the lungs. There is also relative subpleural sparing, which is indicated by red arrows.
User	Are there any devices implanted in this patient?
Language GPT-4	Yes, there is an endotracheal tube implanted in the patient, as indicated by the yellow line in the image.

Table 2: Example comparison of medical visual chat and reasoning capabilities. The language-only GPT-4 is considered as the performance upper bound, as the golden captions and inline mentions are fed into GPT-4 as the context, without requiring the model to understand the raw image.

The Integration of Medicine and AI Presents Diverse Application Possibilities

The combination of medicine and AI is opening up a plethora of possibilities on the application front. From various perspectives such as AI + devices, AI + the entire clinical process, AI + critical clinical decision-making, and AI + lifelong health management, medical AI is forging a new ecosystem for diagnosis and treatment. For example, in coronary CTA diagnosis, following a CT scan, AI can aid in detecting narrowing and plaques while generating structured reports, thereby enhancing the success rate and efficiency of coronary artery reconstruction. Moreover, the AI decision-making track can assess treatment outcomes based on different surgical approaches by simulating digital twin organs, enabling personalized therapy. In the realm of diagnosis and treatment, the integration of electronic health records with AI can improve hospital referrals, information storage, and diagnostic efficiency, thus supporting personalized healthcare.

Furthermore, AI is also empowering the field of innovative drug development, enhancing the efficiency of drug discovery and development. Through machine learning, deep learning, and other AI technologies, pharmaceutical companies can reduce the time and resources required for drug research and improve the success rate of clinical trials. Some companies like Recursion and NVIDIA are actively investing in AI drug development, keeping the global AI pharmaceutical market highly dynamic.

Chunyuan Li
《LLaVA-Med: Training a Large Language-and-Vision Assistant forBiomedicine in One Day》

Applications and Prospects of AI in the Healthcare Industry

Diagnostic Testing/Pathological Diagnosis

Modern wearable devices, such as smartwatches and health bands, can continuously monitor users' physiological data in real-time, such as heart rate, blood pressure, blood sugar, and more. When combined with AI technology, these devices can offer users more precise health recommendations and alerts.

Disease Prevention

Through the analysis of users' physiological data and lifestyle habits, AI technology can predict users' disease risks and provide personalized prevention recommendations.

Rehabilitation Treatment

For individuals with chronic diseases and rehabilitation patients, wearable devices can monitor their recovery progress and provide them with personalized rehabilitation plans based on AI model recommendations.

AI Leading the Future of Healthcare, Making Health Smarter

AI technology has brought unprecedented opportunities to the healthcare industry. From early disease diagnosis and personalized treatment to patient health management and rehabilitation, AI plays an irreplaceable role. However, we should also recognize that the application of AI technology in the medical field faces numerous challenges, such as data privacy and security, technology adoption and acceptance, ethical and legal issues, and more. Only by effectively addressing these challenges can AI technology achieve broader and deeper integration into the healthcare sector.

With the continuous advancement of technology and the ongoing development of the medical industry, we have reason to believe that the healthcare sector of the future will be smarter, more efficient, and more human-centered. As consumers, we can look forward to receiving higher-quality and more personalized healthcare services.

Applications and Prospects of AI in the Healthcare Industry

AI can be employed for highly accurate disease diagnosis, aiding physicians in improving precision and speed.

Diagnostic Testing

AI can be employed for highly accurate disease diagnosis, aiding physicians in improving precision and speed.

AI + Decision-Making: AI models can assist doctors in making more precise decisions. Through digital twin technology, AI simulators of organs can evaluate the effectiveness of different treatment options, thus enabling personalized therapy.

Hospital Information Technology

The demand for data processing in areas such as hospital management, health insurance billing, and quality control is rapidly increasing, leading to rampant growth in medical information technology. Intelligent medical record management, patient information management, and medical resource allocation contribute to enhancing the efficiency of healthcare institutions, shortening diagnosis and treatment processes, and improving healthcare quality.

Medical Imaging Diagnosis

The application of AI technology in medical imaging diagnosis has garnered widespread attention and research. Through deep learning and other advanced algorithms, AI can assist physicians in more accurately identifying and interpreting medical images, such as X-rays, MRIs, and CT scans. For instance, some AI systems can automatically detect abnormalities in the lungs, aiding in the timely diagnosis of lung cancer. In coronary CTA diagnosis, AI can analyze CT scan images, detect narrowing and plaques, and generate structured reports, thereby improving the success rate and efficiency of coronary artery reconstruction.

Disease Prediction and Prevention

AI can analyze vast amounts of patient data to predict the risk of certain diseases. For example, by analyzing a patient's genetics, lifestyle, and other health data, AI can predict the risk of diabetes or heart disease. This predictive capability can help doctors take preventive measures in advance, reducing the occurrence of diseases.

Personalized Treatment

Through AI technology, doctors can create personalized treatment plans for each patient. For instance, based on a patient's genetics and other health data, AI can recommend the most suitable medications and treatment methods. This personalized approach to treatment can improve its effectiveness while reducing the risk of side effects.

Drug Discovery

AI technology also plays a crucial role in drug development. By analyzing vast amounts of biological and chemical data, AI can assist scientists in quickly identifying new drug candidates. Additionally, AI can predict the effectiveness and safety of drugs, accelerating the drug development process. Generative AI technology also holds promise in discovering new drug targets.

Telemedicine

With AI technology, doctors can provide remote medical services to patients. For example, some AI systems can automatically analyze a patient's health data and provide diagnostic recommendations to physicians. This remote medical service can help patients in remote areas access timely medical assistance.

The integration of wearable devices with AI

Interdisciplinary Collaboration

AI technology is facilitating interdisciplinary collaboration between fields such as biology, medicine, and pharmacy, driving overall progress in the healthcare industry.

Integration in the Internet Healthcare Ecosystem

The integration and application of medical AI in the internet healthcare industry chain offer extensive opportunities and potential, encompassing multiple domains including healthcare, pharmaceuticals, and medical insurance, forming a complete "medicine + pharmaceuticals + insurance" closed loop.

Online Consultation and Internet Hospitals

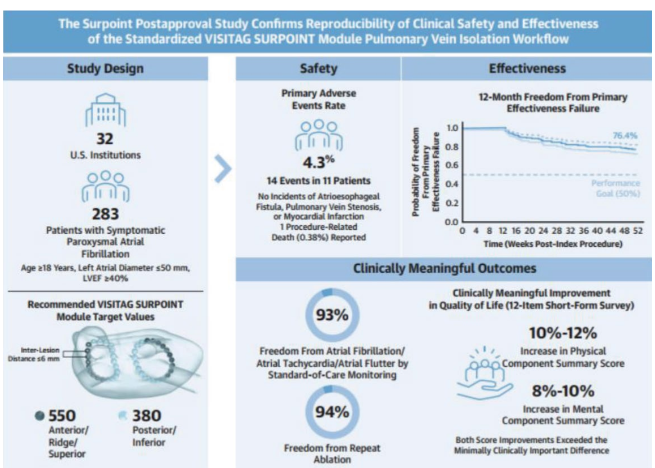
The online consultation market is growing rapidly, and the development of internet hospitals is continually deepening. Patients can access medical advice and diagnoses through internet platforms, thereby enhancing healthcare accessibility and convenience.

AI Healthcare Solutions

Investment in healthcare information technology continues to rise, including investments in contract research organizations (CROs), digital marketing for life science companies, and hospital IT investments. The market for healthcare big data solutions is also growing rapidly, providing comprehensive big data services to the healthcare industry.
It is projected that the penetration rate of healthcare big data solutions will increase from 7.2% in 2019 to 16.2% in 2024.

AI-Powered Ablation

AI-assisted cardiac electrophysiology ablation procedures have demonstrated their effectiveness and safety in several European studies, with the potential to improve the outcomes of atrial fibrillation (AF) ablation. AI can predict lesion depth in radiofrequency ablation procedures, enhancing the effectiveness and safety of ablation.



AI Technology in Smart Homes

With the rapid advancement of technology, artificial intelligence (AI) has become an indispensable part of our lives. From smartphones to self-driving cars, AI applications are ubiquitous. In the realm of smart homes, AI technology brings unprecedented convenience and comfort to our lives. This article delves into the current state of AI tech applied in the smart home industry, as well as future trends and market prospects.

The Current State of the Global Smart Home Market

In terms of Statista, the global smart home market size steadily uplifted from 2017 to 2021, acquiring $104.42 billion in 2021, with a year-on-year spike of 32.56%.

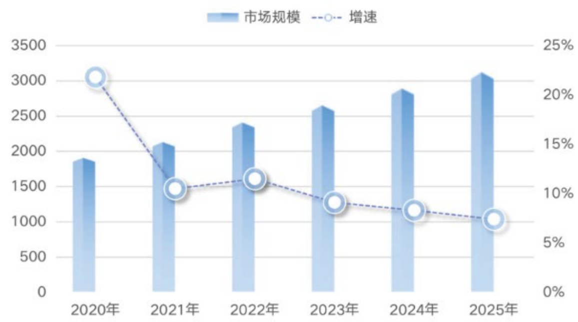
The market size of the U.S. smart home sector has been increasing year by year

The United States, one of the largest and most popular smart home markets globally, has experienced rapid growth in the smart home industry in recent years. Statista indicates that the American smart home market scale increased annually, creating a $28.86 billion value in 2021, ascending by 23.7% from a year ago.

The market size of the smart home sector in China is also continuously expanding.

Pushed by the continuous progress of artificial intelligence and the Internet of Things (IoT), China's smart home market is expanding constantly. China's smart home market grew from ¥260.85 billion in 2016 to ¥514.47 billion in 2020, with an average annual compound growth rate of 18.51%, disclosed "White Paper Of China' s AI+IoH" published by CSHIA Research. Preliminary estimates suggest that in 2021, the Chinese smart home market was approximately ¥580 billion.

It is reported that the Chinese smart home appliance market reached ¥190.74 billion in 2020, and hit ¥213.12 billion in 2021, with a rise of 11.7%. Furthermore, Beijing proposed measures encouraging the introduction of green and smart home appliances to rural areas and promoting trade-in programs for old appliances. The policy is going to stimulate consumer willingness in rural areas, elevate the penetration of smart home appliances in lower-tier markets, and further expand the market. It is estimated that the smart home appliance market in China, by 2025, will accomplish ¥311.95 billion.



The Current Application of AI in Specific Areas of Smart Homes

AI applications in smart home products (excluding household appliances) have become widespread, including smart beds (mattresses), functional sofas, smart curtains, and intelligent toilets.

Smart Beds (Mattresses)	The global market for smart adjustable beds is expanding rapidly, with the Chinese market in the early stages of acceleration. According to data from WiseWise, the global market for smart adjustable beds has expanded rapidly, reaching $5.21 billion in 2022, with a compound annual growth rate (CAGR) of 21.6% from 2015 to 2022. The Chinese smart bed market, as of 2021, had a market size of ¥2.44 billion, with a year-on-year increase of 12.1%, and is still in its initial stages. These beds already offer intelligent elevation, health monitoring, and massage functionalities.
Smart Toilets	AI is primarily used for voice control and health monitoring. The Chinese smart toilet market has experienced rapid growth in recent years, reaching ¥17.7 billion in 2022, with a CAGR of 24.0% from 2017 to 2022.
Functional Sofas	The development of functional sofas is currently in its early stages with limited smart applications. However, in recent years, China's functional sofa penetration rate has rapidly increased from 3.7% in 2016 to 6.2% in 2021. With the assistance of AI, functional sofas' functionality and interaction may see significant enhancements in the future.
Intelligent Curtains	The functionality of current smart curtain products remains relatively limited, with room for further enhancement. In 2021, China's total sales of smart curtains reached ¥12.71 billion, with a Compound Annual Growth Rate (CAGR) of 91.0% from 2016 to 2021. Additionally, smart home appliances have gradually become an integral part of modern life. From backyard equipment to the living room, and kitchen, smart technology endlessly makes our daily lives convenient.

Backyard Equipment

Swimming Pool Water Quality Monitor
This device allows users to test the water quality in their pool through a smartphone application. For example, the WaterGuru Sense can test water quality and adjust water levels.

Smart Lawn Mower
Smart outdoor lawnmowers are meant to map out an entire yard, exemplifying the Worx Landroid S which can automatically mow the lawn and be controlled through a smartphone application.

Outdoor Surveillance Cameras
Outdoor surveillance cameras, like the Arlo Pro 4 Spotlight Camera, use AI to monitor outdoor areas.

Living Room
Smart TVs: Smart TVs are becoming an essential part of every household, integrating Bluetooth and Wi-Fi technologies to provide music streaming services and video-on-demand platforms.

Voice-Controlled Lighting
Synchronizing smart light bulbs and fixtures with platforms like Alexa or Google Assistant allows users to control the lighting in their homes.

Smart Ceiling Fans
When equipped with Bluetooth speakers, smart ceiling fans can help cool a room. The Haiku Home L Series 52-inch smart ceiling fan can be controlled through Bluetooth speakers.

Smart Refrigerators
These appliances can monitor the temperature of the refrigerator and freezer, adjust the filter status, and send alerts in case of system malfunctions.

Smart Ovens
Remote preheating and cooking time monitoring are features of smart ovens.

Smart Dishwashers
These appliances can send notifications, such as when more detergent or rinsing is needed.

Other intelligent Devices

Baby Monitors
AI baby monitors with motion sensors and sound alerts.

Pet Feeders
These can establish automated feeding schedules for pets and come with smart cameras and microphones.

Smart Plugs
Devices that can be inserted into standard outlets, allowing users to control the power of connected devices through applications or voice commands.

Future Analysis and Prospects for AI in the Smart Home Industry

Based on current trends, the integration of artificial intelligence with smart homes is inspiring a global household revolution. Consumer interest in smart homes is increasing, both in China and other regions worldwide, especially in Europe and North America. The market is experiencing rapid growth.

China's smart home market is expanding rapidly. In the long term, the market is expected to grow from ¥3,558.2 billion in 2020 to ¥4,801.2 billion in 2025, with a CAGR of 6.2%. Smart home appliances dominate the market, with a projected market size of ¥3,892.5 billion in 2025.

The Chinese government's "14th Five-Year Plan" emphasizes the development of new smart cities, providing a favorable policy environment for the smart home industry. Local companies like Xiaomi and Huawei have won a slice of market shares of the smart home, and they will likely compete with international giants such as Google and Amazon in the future.

Europe and North America are leaders and major participants in the global smart home market. Data from Statista presented that the U.S. smart home market stood at roughly $40 billion in 2022, and it is predicted to exceed $60 billion by 2025. Europe anticipates a €50 billion sales of smart homes by 2025 in the face of constant growth.

In the Asia-Pacific region, India has significant growth potential in the smart home market. By 2025, the Indian smart home market's size may win $6 billion. The Indian government is implementing initiatives like "Digital India" and "Smart Cities," providing opportunities to enlarge the smart home market. However, data security and privacy concerns have also caught the government's attention and may influence the market.

Given the rapid advancement of AI technology, smart homes have shifted from conceptual to practical, integrated into the daily lives of global consumers. Looking ahead to the future of smart homes.

Closer Integration of Smart Homes with Other Industries: As technology progresses, the integration of smart homes with industries like healthcare, entertainment, education, and more will provide users with comprehensive services.

Emphasis on Data Security and Privacy: As smart home devices handle increasingly more personal data, ensuring data security and user privacy will become a significant challenge.

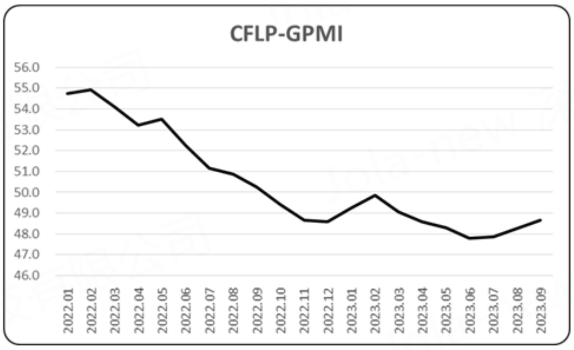
Environmental-Friendly and Energy-Efficient Smart Homes: With growing global environmental concerns, future smart homes will place greater emphasis on environmental sustainability and energy efficiency, with products like smart thermostats and light bulbs moving towards more energy-efficient solutions.

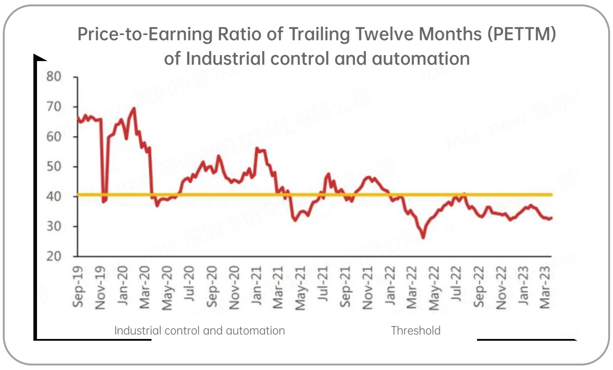
Artificial intelligence (AI) has penetrated various segments of the smart home industry, including smart lighting, security monitoring, energy management, home entertainment systems, and home health and fitness devices. The smart home market is experiencing significant growth in Europe, North America, and the Asia-Pacific region, each with its unique market dynamics and challenges.Europe and North America, as advanced technology markets, have established mature smart home ecosystems, while the Asia-Pacific region, especially China and India, is rapidly emerging as the world's largest smart home market due to its vast consumer base and tech-savvy population.

However, the success of this industry depends not only on technological advancements but also on consumer needs, cultural differences, and price sensitivity. In the future, as AI technology continues to advance and the market matures, we can anticipate a more intelligent, interconnected, and personalized home life. To achieve this vision, industry, government, and consumers must work together to ensure that technological developments genuinely meet people's needs while safeguarding data security and privacy.

Industry automation

Macroeconomic -- Global manufacturing PMI





The development of global industrial automation is closely linked to the Global Manufacturing Purchasing Managers' Index (PMI).

The European manufacturing PMI, although, eventually reversed the decline this month, the index remains at a relatively low level. Downward pressure has eased from August in Europe, but the industry is still trapped in its weak performance. The European Central Bank has raised interest rates to historic highs, caused by the persistent inflationary pressures in Europe. The current interest rate level, whose fluctuation depends on the inflation rate, will be maintained for some time, the European Central Bank disclosed.

The global manufacturing PMI was 48.7%, in September 2023, edging up 0.4% compared to the previous month for the third consecutive monthly increase.

The continuous rise in oil prices has hugely aggravated inflation in South Africa and Egypt, adversely affecting the recovery of their manufacturing sectors. The South African Ministry of Mineral Resources and Energy confirmed that fuel prices in South Africa will highly ascend from October on, shocked by the inflation of international crude oil prices and fluctuations in the ZAR exchange rate.

A report released by the Institute for Supply Management unveils that U.S. manufacturing remains rallying since the third quarter. In September, the U.S. manufacturing PMI was 49%, climbing by 1.4% sequentially with a larger increase, marking the third consecutive monthly expansion. In terms of each index in every aspect of manufacturing, the production and demand witnessed accelerated recovery in the U.S. from the month before. The production seems to recover faster than the demand. The production index went up to above 52% while the new order index rose above 49%, both enlarging by more than 2% compared with August.

Overall, spurred by the rapid rebound in China and the sustained robust growth in India, the trend in Asian manufacturing remains stable.

The industrial control market size shrunk generally, lingered by the dull manufacturing as a whole in the first half of 2023. However, top manufacturers of industrial control still achieved steady revenue growth by seizing structural opportunities in the industry. Based on data from the first half of this year, global PMI is still at low levels, and the overall market stays at a slight decrease until the end of 2023. It is reported that there is a relatively high inventory in the industrial automation market, and distributors are not keen to clear the stocks, being rather prudent over the market demand. Having said that, there are differences within the industry; project-based sectors (such as chemicals, mining, power, and municipal services) are generally promising, while OEM industries (such as printing, plastics, packaging, machine tools, etc.) dramatically slump. Emerging industries (such as new energy and lithium batteries) are growing rapidly due to government policy support.

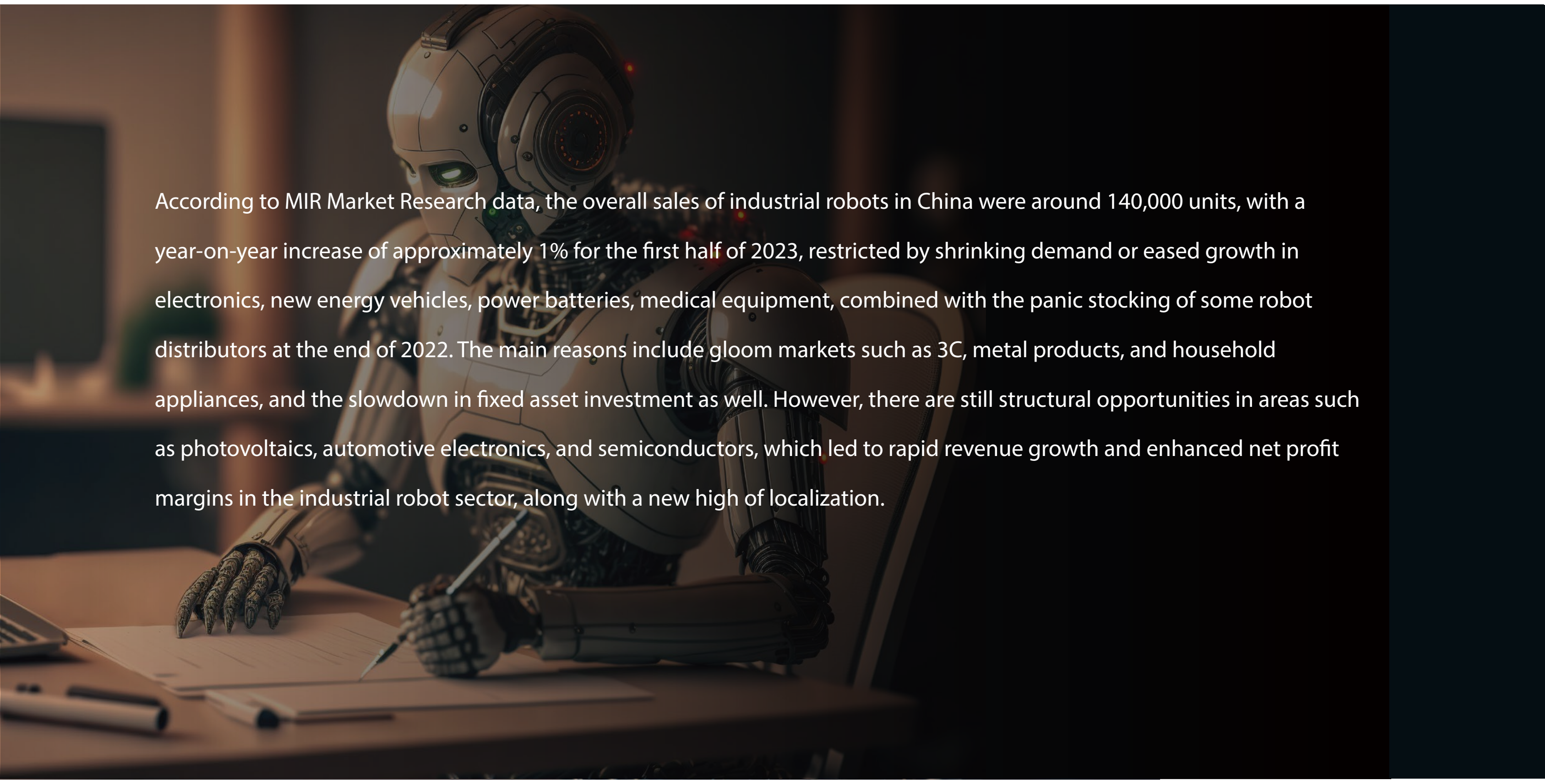
Manufacturing PMI Index for Major Countries and Regions Worldwide in September 2023

Regions	September 2023	August 2023	MoM
USA	49	47.6	1.4
Japan	48.5	49.6	-1.1
Germany	39.6	39.1	0.5
France	44.2	46	-1.8
Italy	46.8	45.4	1.4
Nederland	43.6	45.9	-2.3
Ireland	49.6	50.8	-1.2
Greece	50.3	52.9	-2.6
Spain	47.7	46.5	1.2
Austria	39.6	40.6	-1
England	44.3	43	1.3
Canada	47.5	48	-0.5
Russia	54.5	52.7	1.8
Brazil	49	50.1	-1.1
India	57.5	58.6	-1.1
Indonesia	52.3	53.9	-1.6
South Korea	49.9	48.9	1
Vietnam	49.7	50.5	-0.8
Turkey	49.6	49	0.6
Poland	43.9	43.1	0.8

Industrial automation equipment and systems are the core of industrial automation applications. The main products include human-machine interfaces, controllers, variable frequency drives, Servomechanism, stepper motors, sensors, and related instrumentation. These can be further divided into Programmable Logic Control Systems (PLC), Distributed Control Systems (DCS), Plant Asset Management, Electronic Control Units (ECU), Manufacturing Execution Systems (MES), Product Life Management (PLM), Machine Vision Systems, Human-Machine Interfaces (HMI), Supervisory Control and Data Acquisition (SCADA), Computer Numerical Control (CNC) routers, and more. PLC, servo, and variable frequency drives have seen significant development in recent years. In particular, servo and variable frequency drives accomplish a substantial rise in the market share.

In the second half of this year and the following year, project-based businesses are expected to be more prosperous in contrast to the OEM market, including chemical, power, mining, and others.

Robot
a symbol of industrial automation



According to MIR Market Research data, the overall sales of industrial robots in China were around 140,000 units, with a year-on-year increase of approximately 1% for the first half of 2023, restricted by shrinking demand or eased growth in electronics, new energy vehicles, power batteries, medical equipment, combined with the panic stocking of some robot distributors at the end of 2022. The main reasons include gloom markets such as 3C, metal products, and household appliances, and the slowdown in fixed asset investment as well. However, there are still structural opportunities in areas such as photovoltaics, automotive electronics, and semiconductors, which led to rapid revenue growth and enhanced net profit margins in the industrial robot sector, along with a new high of localization.

The robot industry is experiencing tremendous growth

The latest report from the Association for Advancing Automation (A3) shows that North America, after enjoying record growth of the robot market in the past two years, has seen declines for two consecutive months in the first half of 2023. From April to July 2023. North American companies ordered 7,697 robots worth $457 million, a 37% plummet in robot orders and a 20% decrease in value compared to the same period in 2022.

The surge in orders over the past two years was driven by the automotive industry. Automotive industry customers accounted for two-thirds of the overall orders, but this ratio is gradually approaching a 50/50 split. The demand from the automotive industry continuously creeps up while the construction, agriculture, and healthcare sectors are recognizing the salient applications of robots, which are about to generate substantial growth in the robot industry.

The boom of the manufacturing sector in India, as well as nearshore outsourcing from North America to places like Mexico, boost the demand for industrial robots in these regions. In particular, India presents a higher level of participation in A3 activities and website visits this year compared to any other country. According to a study by Indian think tank Mordor Intelligence, the size of the industrial automation market in India is expected to uplift from $13.23 billion in 2023 to $25.76 billion by 2028.

There are various ways to classify industrial robots.

Industrial robots refer to articulated robots used in the field of industrial production. They are capable of replacing human labor to perform heavy and repetitive tasks while improving the precision and production efficiency of industrial processes. There are various ways to classify industrial robots based on motion coordinate forms, drive methods, and mechanical structure characteristics. They can be categorized into four main types: articulated robots, SCARA robots, collaborative robots, and Delta robots. They are widely used in the automotive manufacturing and 3C electronics industries, with general applications ranging over material handling, assembly, welding, and painting.

The supply chain of industrial robots

Upstream

Materials/ components

Servo Motor
Manufacturers: Yaskawa, Siemens, Schneider Electric, Inovance, Delta, Estun, etc.

Reducer
Manufacturers: Nabtesco, Hermle, Sumitomo Electric Industries, Leaderdrive, LikeDrive, Nantong Zhenkang Welding Electromachinery, etc

Controller
Manufacturers: KEBA, ABB, FANUC, KUKA, Estun, Siasun, HSRobotics, Inovance, etc

Mid-stream

Robots manufacturing

International manufacturers
(exclude China): Fanuc, ABB, Yaskawa, KUKA, Stäubli, Sumitomo Electric Industries, Nachi-Fujikoshi Corp., Epson, Kawasaki Robotics, OCT, etc.

China's manufacturers
Estun, Siasun, Efort, Topstarltd, HSRobotics, Guangzhou CNC Equipment Co., Ltd, STSrobot, Step Electronic, etc.

Downstream

System integration

The manufacturers integrate materials and robots with intelligent management software to build automated production lines for factories.

International manufacturers
(exclude China): Fanuc, Yaskawa, Siemens, Schneider Electric, ABB, Durr, etc

Applications

Manufacturing Plants

Automobiles

Household appliances

Consumer electronics

Metal processing

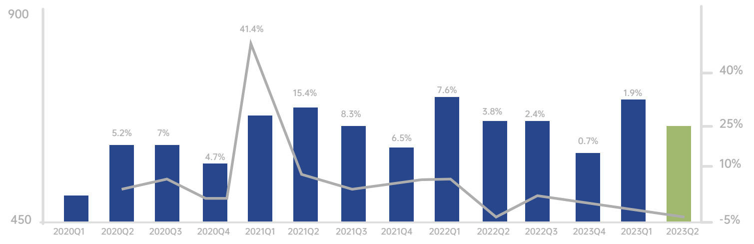
Logistics

Higher application ratios in the automotive and electrical electronics fields.

The key components of industrial robots encompass reducers, servo motors, and controllers, which respectively occupy 35%, 20%, and 15% of the total cost of industrial robots. The end-user market comprises the automotive, household appliance, aerospace, and other industries.

Globally, the automotive and electrical/electronic markets are the main application markets of industrial robots. It is reported that the industry sizes for the two areas gained \$3.545 billion and \$3.472 billion, respectively. In terms of the overall market size, it has experienced slight fluctuations in recent years due to the impact of the pandemic. However, it is expected that the overall market size will steadily grow as the influence of the pandemic diminishes.

Review of China's industrial automation market in the first half of 2023



The overall market size of automation in the first half of 2023 was CNY 151.9 billion **(nearly $20.78 billion)**

a **2.4%** year-on-year decrease.

The OEM market demand has continued its sluggish state for a quarter.

Looking ahead to the second half of 2023

MIR Intelligence believes that the automation market continues to face downward pressure with limited prospects. The following are a few factors contributing to this decline.

- 2023 is the low point of the three-year cycle, with the market downturn potentially causing excessive contraction of confidence among agents.
- Certain cyclical industries, such as real estate and manufacturing, might encounter investment declines and subdued demand.
- Sluggish overseas demand significantly impacts the automation market.

In the first half of 2023, the project-based automation market had a size of CNY 94.4 billion ($12.9 billion), mounting by 1.7% year on year. In the first quarter, its market scale achieved CNY 49 billion, with a year-on-year growth of 1.3%. In the second quarter, the project-based automation market was CNY 45.4 billion ($6.7 billion), up 2.1% from the same period of the previous year.

Considering automation products in project-based industries, the second quarter witnessed growth in power, automotive, mining, chemicals, and municipal and public facilities, while the petroleum & chemical sector experienced a severe drop for two consecutive quarters.

Automation companies' strategies for 2023

Many data can now be processed and analyzed in the cloud, making cloud providers essential for enterprise automation. Companies like Siemens and ABB are starting to collaborate with major cloud companies. For example, Honeywell is integrating its hardware with AWS software to provide services such as digital twinning, MES, IIoT platforms, drones, etc., customized for specific industries like electronics manufacturing, chemical energy, aerospace, and autonomous vehicle industries. ABB is partnering with the German data processing company Celonis to integrate global financial data and optimize various business processes (orders, purchases, payments).

Haier, a major household appliance manufacturer, has plunged $700 million in smart warehousing over the past few years in its nine US factories. They adopted a large number of automated guided vehicles (AGVs), autonomous mobile robots (AMRs), and other equipment to replace high labor costs. These systems are used for picking, packaging, shipping, and other critical processes.

IKEA, the Swedish furniture company, possesses numerous smart warehouses worldwide, featuring a significant number of smart conveyors and autonomous mobile robots. They are also recently adapting drones for scanning barcodes and inventorying goods.

Siemens has announced a €200 million investment in building a new factory in Singapore to expand its industrial automation business and serve the rapidly growing Southeast Asian market.

Outlook

As the workforce ages, younger people are less inclined to work in manufacturing plants. The trend is moving towards using automation to replace manual labor. While governments do not give any subsidies, the demand from plants is ascending. The new energy sector is expected to continue to thrive, including full-vehicle manufacturing and power management. In addition, there are specific opportunities in energy management, such as the intelligence and automation of water meters, ammeters, and gas gauges, or using variable frequency equipment, and equipment upgrades to realize the energy-save.

The Battle for AI Chips **Who Will Reign Supreme** in the Competition?

Increasingly, artificial intelligence (AI) computing makes its way into areas like aiding robot learning, training large language models (LLMs), and AI inference, coupled with autonomous driving, smart healthcare, and factory automation progressively integrating AI technology. Therefore, the demand for AI chips is meant to continue its robust growth.

Global hardware sales for AI are projected to gain $53.4 billion in 2023, a YoY growth of 20.9%, According to the latest report from market research firm Gartner. The company added that AI semiconductors will maintain double-digit growth, with revenue soaring by 25.6% to reach $67.1 billion by 2024. Revenue of global AI chips is anticipated to hit $119.4 billion in 2027, more than doubling the market size from 2023.

Currently, artificial intelligence technology has permeated various sectors, covering manufacturing, healthcare, transportation, finance, education, and security. Computation power, algorithms, and data represent the three critical elements of AI development, with computing primarily realized by AI chips that have become a battleground for major semiconductor manufacturers.

NVIDIA

NVIDIA seems to be dominating the AI chip arena. It is reported that NVIDIA holds an 82% share of the global AI accelerator market for data centers and monopolizes the global AI training market with a 95% market share, making it the leader in computational power chips of AI large models.

Currently, NVIDIA's AI chips are essentially featured by H100 and A100. According to Omdia, in the second quarter of 2023, NVIDIA shipped 900 tons of H100 AI GPUs, equivalent to over 300,000 H100 units. Omdia revealed that NVIDIA is expected to sell approximately 3,600 tons of H100 GPUs in 2023, roughly 1.2 million units.

Moreover, generative AI models of OpenAI, Google, Meta, Baidu, Tencent, and Alibaba, heavily depend on NVIDIA's AI chips for training. In 2024, the annual growth is expected to stand at 64% in the face of the gradual increase in TSMC's CoWoS production capacity.

In addition to the popular A100 and H100, NVIDIA recently launched the next-generation GH200 Grace Hopper super chip platform designed specifically for generative AI. The new GH200 boasts a rise memory capacity of 3.5 times and bandwidth that is three times greater than the previous version, which means that a server with GH200 possesses 144 Arm Neoverse cores, 282GB HBM3e memory, and 8 petaflops of AI computing power. Production of GH200 is expected to begin in the second quarter of 2024.

AMD

AI as a key priority for AMD has remarkably contributed to its well-performance in the second quarter. AMD' s financial reports disclosed that demand for AI solutions is mounting for data centers. The meetings and communications over artificial intelligence between AMD and its customers are ramping up sevenfold from April to June, with multiple customers initiating or expanding plans for future large-scale deployment of the Instinct accelerator.

AMD has been active in the field of AI since January 2023. At the Consumer Electronics Show (CES) in January 2023, AMD's CEO, Lisa Su, officially introduced the next-generation data center APU (accelerated processing unit), Instinct MI300. Its AI computing and Performance-per-watt elevates eight times and five times compared to the previous MI250. AMD intends to start mass production and supply in the second half of 2023.

In June, AMD announced new AI acceleration chips, Instinct MI300X, and Instinct MI300A, designed for generative AI. These chips can run large models with up to 800 billion parameters on a single chip. They are expected to be shipped by the end of this year. MI300X is positioned as the "second choice" after NVIDIA's H100.

On September 18, 2023, AMD introduced the new EPYC 8004 series processors, enhancing the fourth-generation AMD EPYC CPU family for workload-optimized processors. This caters to intelligent edge applications in retail, manufacturing, and telecommunications, as well as cloud services, storage, and data center applications.

Additionally, AMD showcased a powerful software ecosystem for data center accelerators called ROCm, bringing an open, mature, and complete AI software platform to the market.

Intel

In the fierce competition for AI supremacy, Intel has made significant strides in the AI domain and integrated AI into various hardware products. Intel's AI chip offerings consist of three main categories: the specialized Gaudi series for AI, the Xeon Scalable chip series, and GPU product lines.

In July, Intel introduced the Intel Gaudi 2 AI chip, designed for training large language models and built using a 7nm process. With 24 tensor processor cores, Intel is aiming to lead the AI computational power market.

The next-generation Gaudi 3 will use a 5nm process and offer improved performance, including a 4x increase in BF16 performance, a 2x increase in computational performance, a 1.5x increase in network bandwidth, and a 1.5x increase in HBM capacity.

In addition, Intel's PC and server chip roadmaps now more deeply integrate AI capabilities. Intel will launch the Core Ultra processor, codenamed Meteor Lake, on December 14. This processor will include Intel's first integrated neural network processor (NPU), delivering efficient AI acceleration and on-device inference experiences for PCs.

So, in the race for AI chip supremacy, it remains to be seen which company will ultimately lead the way and define the future of AI computing. The competition is fierce, with NVIDIA, AMD, and Intel making significant strides to dominate this crucial sector.

Broadcom

Broadcom is the world's second-largest AI chip company, following closely behind NVIDIA. According to its latest financial reports, its semiconductor business is predominantly driven by AI. Major customers for Broadcom's AI chip business include Google, Meta, and Microsoft. Google's self-developed AI chip, the TPU Tensor Processor, is produced by Broadcom. A report released by JPMorgan Chase analyst Harlan Sur at the end of May stated that Google's TPU orders could generate over $3 billion in revenue for Broadcom. Meta is also using AI chips developed by Broadcom itself, though Meta purchased a small amount of AI chips.

At the 49th European Conference on Optical Communication (ECOC 2023), Broadcom introduced its new 5nm 200 Gbps per channel Pulse Amplitude Modulation (PAM) digital signal processor (DSP) named Sian BCM85822, which features integrated laser drivers. It operates at speeds of 200 Gbps per channel using PAM-4 modulation and began sample shipments in August, with plans for an official launch in mid-2024. It is primarily intended for 1.6T deployments in traditional data center networks but lays the foundation for the next-generation 1.6T/3.2T market in AI clusters.

Microsoft

As AI technology continues to advance, more and more companies are focusing on the AI hardware sector. Microsoft is planning to launch its first AI chip, called "Athena," in November. The upcoming Athena chip from Microsoft is similar to NVIDIA graphics processing units (GPUs) but designed specifically for training and running large language models on servers. Moreover, the chip was developed in collaboration with OpenAI, and it has been secretly tested by OpenAI. It is designed for training and running large models and has shown excellent performance, at least on par with mainstream chips from companies like Amazon and Google.

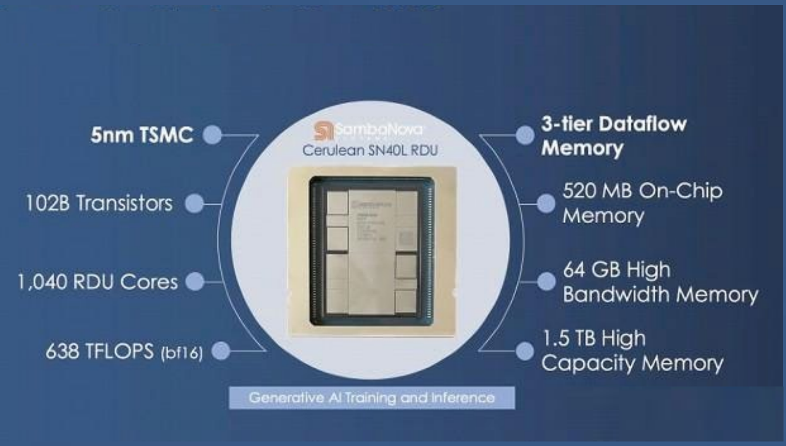
Samsung

Samsung held a Technology Day event on October 5th in Silicon Valley, sharing its recent collaborations in AI chip development. They aim to manufacture chips for leading-edge 4nm designs for fabless IC design companies. The president of Samsung Electronics, Park Yong-in, mentioned that generative AI has become the biggest trend this year, necessitating the most powerful foundational technology to handle data and make AI available. Samsung is actively paving the way for the new era of generative AI.

Samsung will collaborate with fabless companies to develop a new AI chip named Rebel, using 4nm process technology and packaging the most advanced HBM3E memory chips. Although TSMC currently dominates global AI chip manufacturing, Samsung's wafer foundries have recently received orders from AI chip startups in Korea, Canada, and the United States. On October 2nd, Canadian company Tenstorrent claimed that its next-generation AI chips will be produced by Samsung's foundry in the United States.

SambaNova

SambaNova is a rising star in the AI chip field and is considered one of the top 10 unicorns in the United States. SambaNova recently launched the new AI chip SN40L, manufactured using TSMC's 5nm process, containing 102 billion transistors, and achieving a peak speed of 638 teraFLOPS and 1.5 terabytes of memory, supporting sequence lengths of up to 256,000 tokens. The SN40L chip is designed to run models over twice the size of OpenAI's ChatGPT advanced version.



SambaNova's chips and systems have already attracted significant clients, including leading supercomputer labs worldwide, such as Japan's Supercomputer Fugaku, the U.S.'s Argonne National Laboratory, Lawrence Berkeley National Laboratory, and consulting firm Accenture.

Furthermore, tech giants like Google, Amazon, and Tesla have ventured into self-developed AI chips. Amazon has two AI-specific chips, Trainium and Inferentia. Google not only upgraded its TPU from the version in 2016 to TPU v4 in 2020 but also secured orders from prominent entities like OpenAI and Midjourney. Tesla, the electric car manufacturer, has rolled out two self-developed chips, full self-driving (FSD) and the Dojo supercomputer. FSD chips are used in Tesla vehicles for autonomous driving, while Dojo D1 chips power the Tesla Dojo supercomputer.

Moreover, as the United States intensifies strict restrictions on high-performance chip exports, China's AI chip development is on the rise. Tech companies in China like Huawei, Alibaba, Baidu's Kunlun, Biren tech, Cambricon, ILuvatar CoreX, and Vastai tech are actively competing in the GPU arena, achieving notable results.

In this competition for AI chips, NVIDIA remains the undisputed "AI computational power champion," with the A100 and H100 chip series at the pinnacle. Meanwhile, these major chip manufacturers are also joining the market race due to the substantial demand, even though self-developed AI chips come with substantial financial and time costs. In the age of AI, being at the forefront is an imperative strategic move.

R&A Electronics

Since our inception in 2015, we have been diligently developing our channels and leveraging our business intelligence and an expansive supplier network to procure inventory directly from toptier suppliers. Our objective is simple: offering world-class procurement services to global OEM/EMS/ODM partners at a highly competitive pricing. R&A has positioned ourself as one of the most rapidly growing semiconductor distributors in the world.

IC The World

IC The World is a showcase for the global economy and semiconductor industry. We dissect the global economy and chip market based on all relevant data. This issue of the journal aims at the trending sector--Artificial intelligence, presenting the applications of AI across various areas and the significant influence it has on the whole market upgrade. We delve into diverse aspects of the semiconductor market, including technological trends, market share, and competitive landscape.

IC The World

R&A electronics

www.randa.sg

Singapore

Address | Suntec Tower 5，Temasek Blvd #17-06 Singapore 038985
Tel：+65 86676866

HongKong

Address：20/F TML Tower, 3 Hoi Shing Road, Tsuen Wan, NT, Hong Kong
Tel：852-27111328

Shenzhen

Address：24 F, Tower B, Chuangyi Technology Building, No. 198, Keji Zhongyi Road, Nanshan District, Shenzhen, Guangdong
Tel：86-755-83170380

Shanghai

Address：Room 2303, Tower B, SOHO Zhongshan Plaza, No. 1065 West Zhongshan Road, Changning District, Shanghai

Shijiazhuang

Address：Room 1749-1750, Tower B, Happy Land 700, Zhongshan West Road, Qiaoxi District, Shijiazhuang City, Hebei Province