

Metaverse

The New Economic Road in Asia-Pacific

162212

www.freemove.com is a for-fee contents mall that provides the world's best contents services. designers. freemove has a design contents library that enables.

162212

www.freemove.com is a for-fee contents mall that provides the world's best contents services. designers. freemove has a design contents library that enables.

120101

www.freemove.com is a for-fee contents mall that provides the world's best contents services. designers. freemove has a design contents library that enables.

Produced by SenseTime Intelligent Industry Research Institute

[About the Tech4SDG]

The Technology for Sustainable Development Goals Alliance for Asia (Tech4SDG for short) is committed to many aspects such as increasing engagement in society, facilitating industry-university-research (IUR) interaction, conducting exchanges and communications in science and technology and studying its ethics. It also deeply involves in the development of standards, the compilation of industrial case studies, and the publication of the subsequent results to promote sustainable development in Asia. The alliance is composed primarily of technology enterprises, research institutions, think tanks, universities, experts, and scholars in Asia.

[About SenseTime Intelligent Industry Research Institute]

Relying on the artificial intelligence technology of SenseTime's massive data and supercomputing capabilities, it deeply cultivates AI industry practice and cutting-edge research, and participates in reports and researches at the Ministry of Science and Technology, Ministry of Industry and Information Technology, National Development and Reform Commission and other ministries and commissions based on case results; Influential AI industry frontier think tank.

For more information about us or to obtain the electronic version of this report, please visit the website www.sensetime.com/cn or follow our WeChat Official account:



Contact Information: yangyan1@sensetime.com

Contents

| | |
|---|-----------|
| Foreword | 1 |
| Conception | 4 |
| Conception from a Technological Perspective: The Essence of the Metaverse | 5 |
| Three Transitions from Traditional Internet to the Metaverse | 6 |
| Transition 1: The transition in media access delivers a more immersive and natural HCI/HMI experience | 7 |
| Transition 2: The conceptualization of digital natives (user transition) determines the future of the metaverse | 8 |
| Transition 3: The Web 3.0 definition of data rights is likely to alter how platforms create value | 9 |
| An Illustration of the Metaverse | 10 |
| Creation | 12 |
| Three Forces Integrating Virtuality with Reality that Create the Metaverse | 13 |
| Creativity: Accelerating the creation of the virtual world | 13 |
| Connectivity: Realizing the connection between digital and physical worlds | 13 |
| Integration: Pushing the integration of virtuality and reality and intelligent development of the metaverse | 14 |
| The Three Infrastructures that Consolidate Metaverse Productivity | 14 |
| Engine: A low-threshold and cross-terminal creation environment | 14 |
| Algorithm: Accelerating "creation," achieving "connection" and promoting "integration" | 15 |
| Computing power: Supporting the massive computation of the metaverse | 15 |
| Empowerment | 16 |
| SenseMARS Mixed Reality Platform: Engine Infrastructure that Creates and Designs the Metaverse | 17 |
| SenseMARS Avatar: "We/us" in the metaverse | 18 |
| SenseMARS Agent: "They/them" in the metaverse | 19 |
| SenseMARS Reconstruction: Digital reconstruction of the physical world | 19 |
| SenseCore Universal AI Platform: Supporting the algorithms and computing power infrastructure of the metaverse | 20 |
| Cases | 22 |
| Case 1: Asia's "unmanned" help desk created by AEON | 23 |
| Case 2: Zepeto - Custom avatars through rapid face molding | 25 |

| | |
|--|-----------|
| Case 3: AR Navigation at Suvarnabhumi Airport, Thailand | 25 |
| Case 4: Watching games via AR interaction in a Japanese baseball stadium | 27 |
| Case 5: Riyadh Season's immersive AR journey | 28 |
| Case 6: Schwarzkopf - AR hair-dyeing trial | 31 |
| Case 7: AR Digital cultural and creative platform | 32 |
| Epilogue | 33 |

Foreword

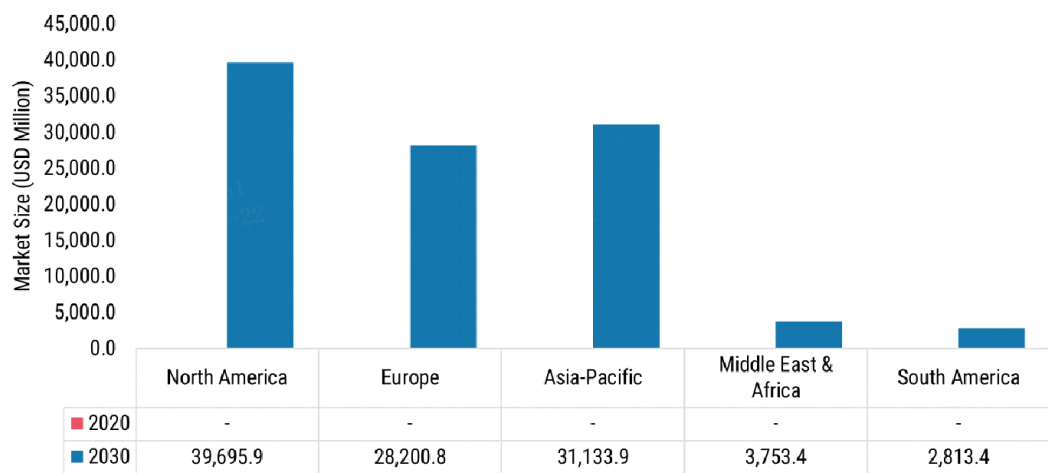
If 2021 was the year in which the "metaverse" became a widely known and understood concept, 2022 is the year that the notion of the "metaverse" will receive recognition and affirmation. In just two years, stakeholders from all over the world, including governments, venture capitalists and tech giants, have spontaneously flocked to the idea of the metaverse, setting off a spectacular round of competition.

| Region | 2021 | 2022 | 2023 | 2024 | 2025 |
|----------------------|-----------------|-----------------|-----------------|-----------------|-----------------|
| North America | 11,055.2 | 12,544.0 | 14,741.2 | 16,785.4 | 17,892.1 |
| Europe | 32.2 | 125.7 | 350.5 | 1,235.3 | 1,774.9 |
| Asia-Pacific | 78.8 | 351.3 | 1,547.3 | 3,514.3 | 4,165.9 |
| Middle East & Africa | 5.7 | 15.6 | 35.2 | 65.9 | 89.5 |
| South America | - | 4.2 | 13.7 | 38.5 | 54.3 |
| Total | 11,171.9 | 13,040.7 | 16,687.8 | 21,639.4 | 23,976.7 |

Source: Company Websites, Annual Reports, Secondary Research, Press Releases, Paid Databases, Expert Interviews, White Papers, Journals, Case Studies, and MRFR Analysis

Figure: Global Metaverse Investments, by Region, 2021-2025 (USD Million)

Asia-Pacific countries, in particular, have given increased attention to their metaverse-related industry development planning. Countries like China, India, Japan, Korea, Singapore, Malaysia, Thailand, and Vietnam are actively promoting the metaverse as an emerging economy within their territories. Merging into the fast lane in 2022, they have increased their investment from USD 78 million in 2021 to USD 351 million. According to estimates by McKinsey & Company, by 2025, metaverse investments will reach USD 4.165 billion in the Asia-Pacific region, accounting for 22.5% of total global investments, with an investment growth rate that will rise by 128%, far exceeding the global growth rate.



Source: Company Websites, Annual Reports, Secondary Research, Press Releases, Paid Databases, Expert Interviews, White Papers, Journals, Case Studies, and MRFR Analysis

Figure: Global Metaverse Market, by Region, 2020 VS 2030 (USD Million)

The Asia-Pacific region accounts for 60% of the world's population, 47.4% of global GDP and 52% of global technology growth, so the metaverse, which benefits from its users' affinity for emerging technologies, has huge market potential to develop in the region. According to the estimates by McKinsey & Company, in the Asia-Pacific region, the metaverse market will reach USD 31.13 billion by 2030 at a compound growth rate of 62.2%. A metaverse report released by the Analysis Group also

indicated that metaverse technology will contribute USD 3 trillion to global GDP over the next ten years, a third of which (i.e., USD 1 trillion) will come from the Asia-Pacific region. In other words, over the next ten years, every US dollar invested in the metaverse industry will bring about an economic growth of USD 3.16, a return on investment of over 300%.

So, what is the metaverse? How is a metaverse world created? As an AI software company, how should SenseTime empower the creation and development of the metaverse, both in the Asia-Pacific region and throughout the whole world? The White Paper "Metaverse: The New Economic Road in Asia-Pacific" centers around four chapters encompassing the conception, the creation, and the empowerment of the metaverse, and specific cases to give all a glimpse into this newly emerging cyber world.

Conception

| | |
|---|----|
| Conception from a Technological Perspective: The Essence of the Metaverse | 5 |
| Three Transitions from Traditional Internet to the Metaverse | 6 |
| Transition 1: The transition in media access delivers a more immersive and natural HCI/HMI experience | 7 |
| Transition 2: The conceptualization of digital natives (user transition) determines the future of the metaverse | 8 |
| Transition 3: The Web 3.0 definition of data rights is likely to alter how platforms create value | 9 |
| An Illustration of the Metaverse | 10 |

Conception from a Technological Perspective: The Essence of the Metaverse

What in the world is "the metaverse"? Literally, the term metaverse is a portmanteau of "meta" (meaning "beyond") and "universe". The metaverse, in effect, is a parallel world that excels beyond reality and is established on the basis of real world. The American, online gaming company Roblox once used eight keywords to describe the main features of the metaverse, which respectively were: "Anywhere, Immersive, Low Friction, Variety, Identity, Friends, Economy, Civility". On the basis of these eight features, we can see that the metaverse, under the Roblox description, should be a parallel and lasting virtual world, where people are able to enter that world whenever and wherever they wish through a virtual avatar, and enjoy its highly-immersive contents and experiences, living in and carrying out social interactions in this world, and it should have an established and completely functional social and economic system.

In addition to describing the metaverse in terms of its key features, we can further conceptualize the essence of the metaverse from a technological perspective. The metaverse, we think, from the perspective of technology development, is the next iterative evolution of networks, led by telecommunications networks, computing/storage, interactive terminals and other IT infrastructures, and follows the development of the mobile Internet.

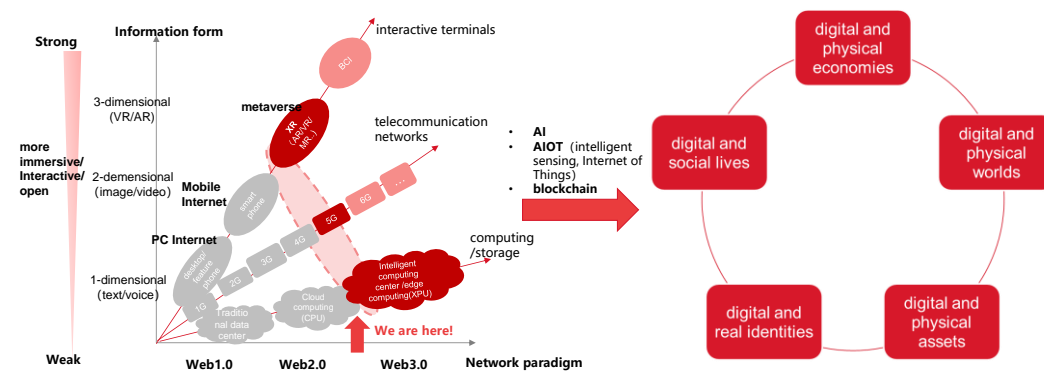
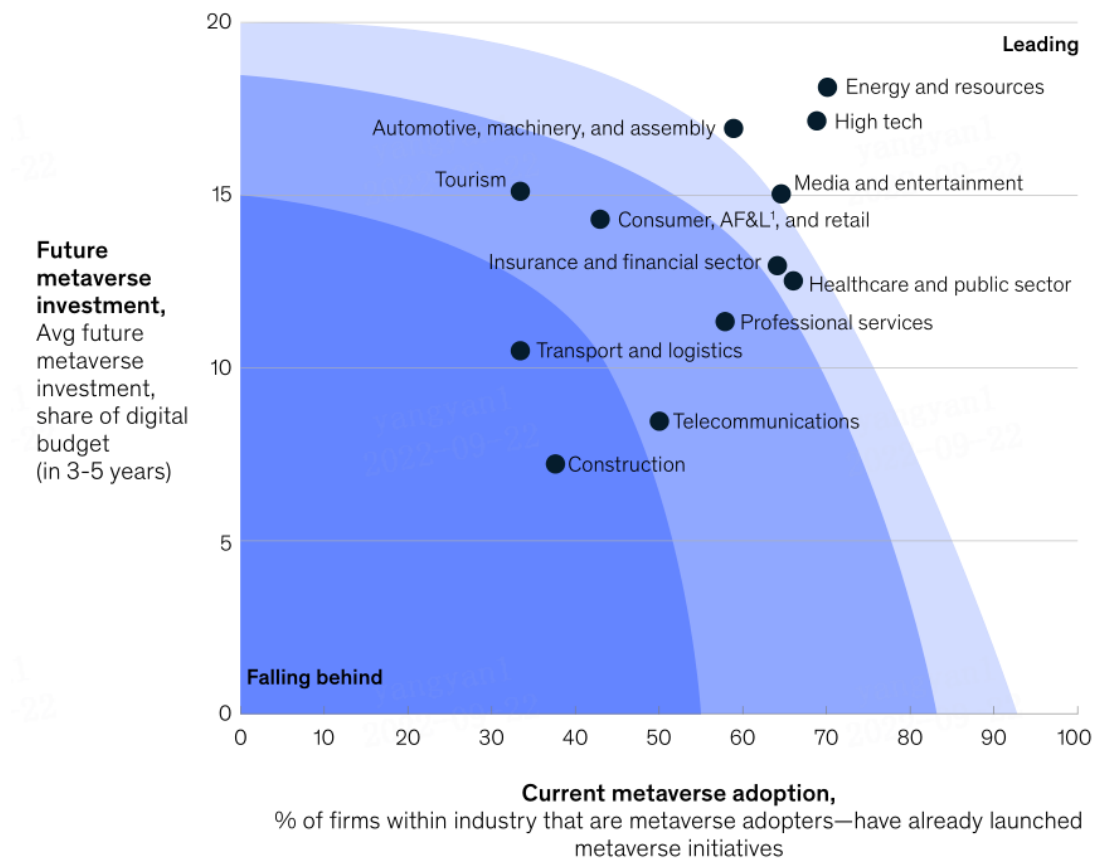


Figure: Metaverse is the next iterative evolution of networks, led by IT infrastructures

As the 5G communication network develops, its higher bandwidth and lower latency will allow us to transmit information and data across more dimensions and with higher throughput, and thanks to the massive scale and energy efficiency improvements in computing/storage infrastructure, such as intelligent computing centers and edge computing, we can more efficiently store and compute large volumes of complex information and data. Given the popularization and application of terminal devices such as AR and VR in daily life and economic production, we have good reason to believe that a new round of network advances will deliver a new experience in the digital world that is more immersive, more interactive, and more open. Additionally, with the application and intensive integration of technologies such as artificial intelligence (AI), AIOT (intelligent sensing, Internet of Things) and

blockchain, a bridge is forming between the digital and physical worlds which will further extend the new socio-economic model characterized by the coexistence and integration of virtuality with reality, thus driving the integration and unification of the digital and physical worlds, digital and physical economies, digital and real identities, digital and social lives, and digital and physical assets.

The metaverse conceptualized from a technological perspective is not simply a parallel "dimensional" universe, but a new "trans-dimensional" world that interacts and integrates with the real world.



¹Apparel, footwear, and luxury.
Source: McKinsey & Company Senior Executive Survey, April 2022

Figure: Sectors leading metaverse adoption today also plan to dedicate a significant share of their digital investment budgets to the metaverse.

According to research conducted by McKinsey & Company, all sectors have started to embrace the metaverse as it further connects to the real world. Over the next three to five years, an increasing number of sectors will dedicate a certain share of their digital investment budgets to the metaverse, and sectors such as energy & resources, automotive, machinery & assembly, technology, tourism, and media and entertainment will become the value creation frontrunners of the metaverse.

Three Transitions from Traditional Internet to the Metaverse

In the move from traditional Internet rules to the new rules of the metaverse that integrate virtuality with reality, purely through our experience in tech development,

we have deduced three predictable transitions, i.e., media, user, and network-paradigm transitions. These transitions may just be the tip of the iceberg in metaverse development, but this iceberg hides innumerable unpredictable changes that are waiting for the spark of human innovation to trigger them. These are the wonderful surprises and hope the metaverse inspires for the development of the whole of human society.

Transition 1: The transition in media access delivers a more immersive and natural HCI/HMI experience

| Revolution | Terminals/ Media | Year | network effect | Informa- tion delay | Immersion | | | | | | Human-Computer Interaction | |
|--------------------------|---------------------|-----------------|-------------------|---------------------------|--------------|-------------|----------|-------|-------|--------------------------|-------------------------------|--|
| | | | | | Hear- ing | Vis- ion | Touch | Smell | Taste | Information dimension | | |
| Gutenberg | books | 15th century | 1: 1 | month | | v | | | | | 1D | |
| Age of Electricity | telegraph | 1840 | 1: 1 | day | | v | | | | | 1D | indirect, two-way |
| | telephone | 1880 | 1: 1 | real time | v | | | | | | 1D | indirect, two-way |
| | broadcast | 1920 | 1: N | real time | v | | | | | | 1D | indirect, one-way |
| | movie | 1910 | 1: N | real time | v | v | | | | | 2D | indirect, one-way |
| | television | 1950 | 1: N | real time | v | v | | | | | 2D | indirect, one-way |
| Age of Digitalization | PC | 1990 | M:N | real time | v | v | | | | | 2D | indirect, two-ways (mouse/keyboard) |
| | smart phone | 2010 | M:N | real time | v | v | | | | | 2D | indirect, two-ways (touch screen) |
| | XR | 2020 | M:N | real time | v | v | v | | | | 3D | nature, multidirectional (body movement) |
| | BCI | ? | M:N | real time | v | v | v | v | v | | multidimen- sional | nature, multidirectional (brain waves) |

Figure: The transition in media access delivers a more immersive and natural HCI experience

The transition in new media access enables the further development of human-sense digitalization. As Marshall McLuhan once said, each generation of media upgrade is an extension of human senses. With the transition from traditional PCs to smartphones, and then to smart wearables like AR/VR glasses and haptic gloves, as media iterates and integrates developments in digital technology, people's vision, hearing, touch, and even senses of smell and taste are gradually being simulated in a digital manner. This allows us to acquire and enjoy sensory feelings and immersive experiences in the digital world that are nearly identical to the real world.

HCI/HMI (Human-Computer Interface/Human-Machine Interface) also becomes increasingly more immediate and natural. In the past, we had to browse web pages with a keyboard and a mouse, and then we touched screens to switch between mobile apps. Today, by installing micro sensors or cameras in smart wearables, we can use the blinking of our eyes or changes in facial expressions or gestures to move around the virtual world. In the future, with the development of BCI (Brain-Computer Interface) technology, we will even be able to directly interact with computers/machines using our thoughts. As such, the transition of media is supported by the development of digital technology and upgrades the experience on the user's end.

Transition 2: The conceptualization of digital natives (user transition) determines the future of the metaverse

In the future, the major group dominating metaverse development will be young people that grow up with relevant metaverse technologies, or more specifically, the groups that represent the vibrant force of consumption—young people who comprise Gen Z and the younger demographic of the burgeoning Gen Alpha.

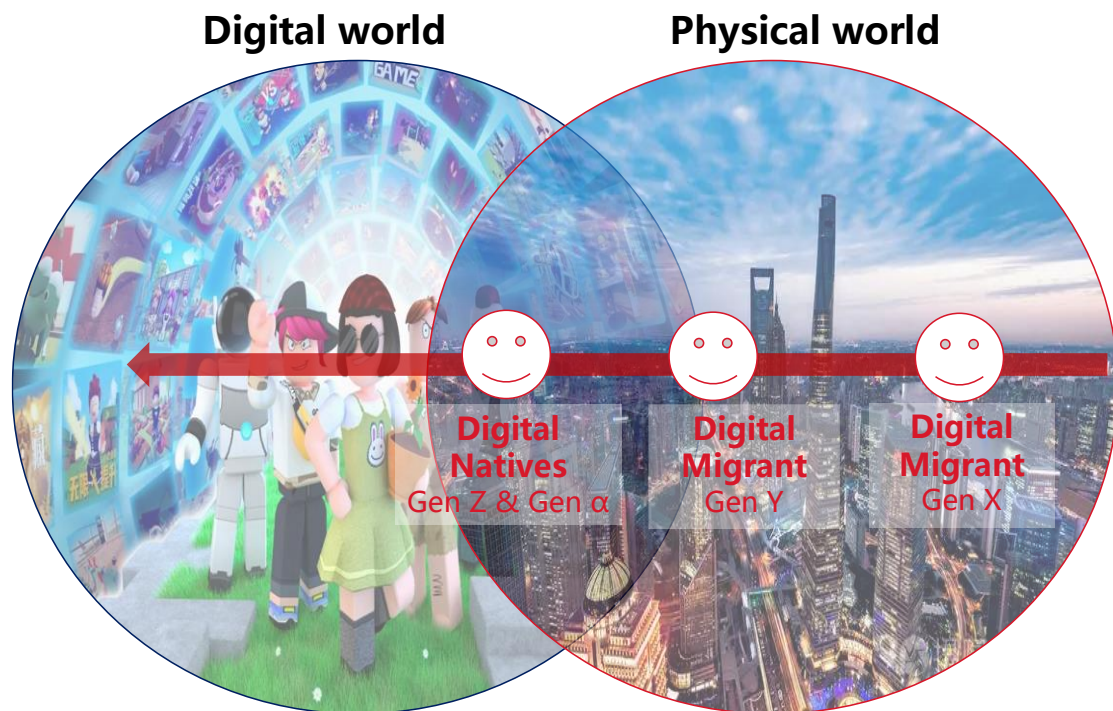


Figure: The global outlook of digital natives is a natural unification that integrates virtuality and reality.

We collectively call these two groups "digital natives". Due to the fact that they have been living in a digital world since they were born, and their communications, interactions, and most of their life are based on the digital world, the global outlook of digital natives is different from other groups (also known as "digital immigrants", Gen Y, Gen X, etc.), and is a natural unification that integrates virtuality and reality. They prefer a mix of both real and virtual consumer experiences. According to a Ypulse study, Gen Z, compared with Gen Y, like creating avatars, enjoy meeting with their friends in a gaming environment, and are more willing to buy virtual goods.

Since digital natives prefer spending time in the virtual world, this has led some commercial brands in the real world, as we have seen, to cooperate with metaverse platforms and provide corresponding virtual products and services based on the consumption requirements of these young people, and also to relocate their youth-oriented marketing sites into the metaverse. They are constantly creating new means of marketing and connecting with digital natives to plan for the future. For example, concerts held in Fortnite; co-branded virtual costumes released in Fortnite by high-end brands like Balenciaga and others; BVLGARI, who, in cooperation with the "Gen Z Arcade", created the special virtual world "BVLGARI ZEPETO World" on ZEPETO, and so on. As we can see, the requirements of digital natives have a direct

impact on the transition of business activities in reality, thus driving the evolution of new business models. Ultimately, these will overturn the value system established by the traditional Internet.

Transition 3: The Web 3.0 definition of data rights is likely to alter how platforms create value

With the application and development of blockchain technology, the network-paradigm of the metaverse will finally evolve to Web 3.0. Compared with the read-only Web 1.0 and the writable and interactable Web 2.0, the biggest characteristic of Web 3.0 is decentralization. We can use a public ledger to store, read, and write data, but this ledger is not controlled or owned by any centralized entity. The data is distributed to and stored at multiple nodes. Errors that occur at any node will have no impact on the data records at the other nodes, so the data is unlikely to be falsified or deleted. At the same time, all of our activities on the blockchain can be recorded and reviewed, so that, in principle, our data rights (including ownership, right of use, revenue rights, etc.) can be identified. Compared with the monopoly held by centralized platforms, which is derived from the failure to identify the boundaries of data rights under the Web 2.0 paradigm, Web 3.0 will change the underlying logic and disrupt the business models of metaverse platforms moving forward. In other words, if a platform wants to create value in the future, it must first clearly define its data rights and income distribution. In addition, gaining income from data monopolies will be difficult, so platforms will need to further open up in order to connect with more users and create higher value for their users.

An Illustration of the Metaverse

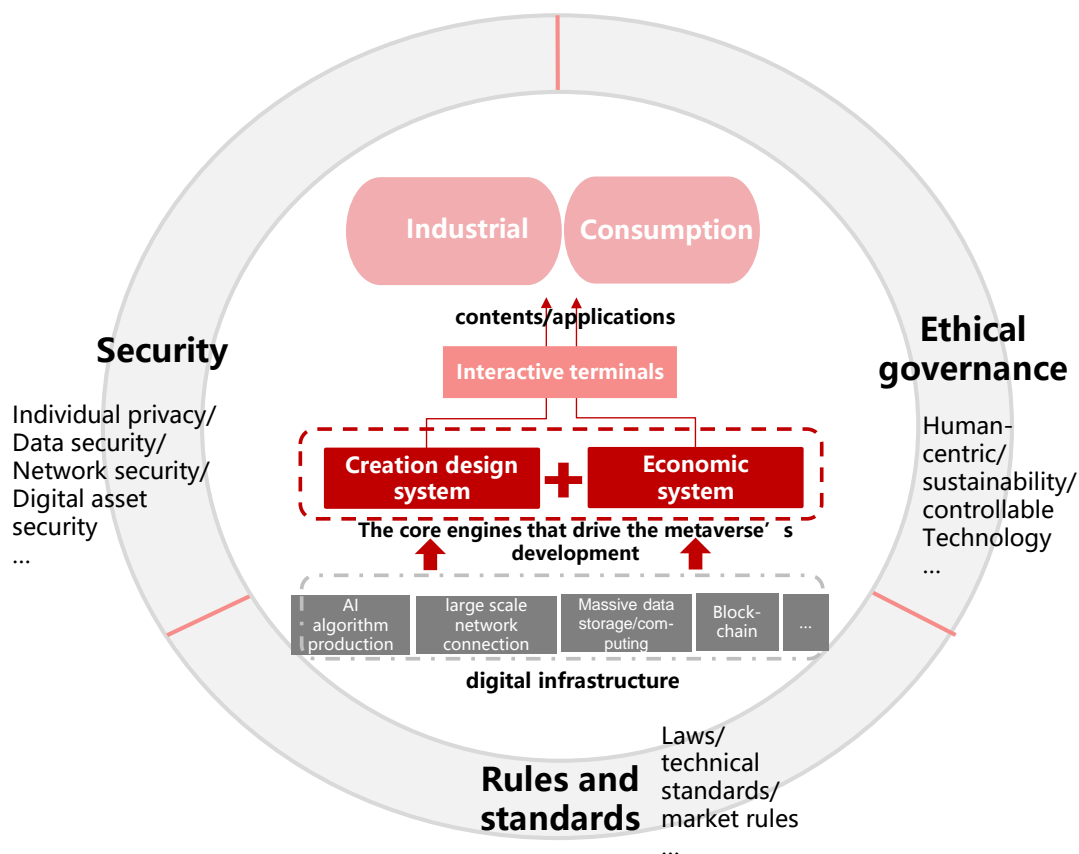


Figure: a simplified chart to provide a general description of what the metaverse looks like

Based on the conceptualization of the metaverse and the difference between the metaverse and traditional networks, we use a simplified chart to provide a general description of what the metaverse looks like.

First, the construction of the digital infrastructure is the foundation of the development of the metaverse. To make the metaverse "immersive, low-friction, and anywhere," an expansive amount of work from both network transmission and storage/computing is required; efficient AI algorithm production can accelerate content production and the distribution process, which in turn greatly enriches the content ecosystem of the metaverse; blockchain supports the efficient operation of the economic system, and ensures the security of digital assets and IDs, thereby guaranteeing the value exchange between digital assets and the transparent implementation of system rules.

Second, based on the digital infrastructure, the core engines that drive the metaverse's development are the creation design system and economic system. Based on the former, people design and create the metaverse world around the concepts of "people, things, and environment," and continue to add rich and diverse digital content, enjoy experiences through interaction terminals, and ultimately create value in both the consumption and production sectors. The latter, a well-functioning economic system, also helps achieve value exchange (making the pie bigger) and value distribution (dividing the pie properly) in the metaverse ecosystem, thus forming a value

"flywheel" covering everything from the production to the application of content, and allows the metaverse ecosystem to develop and thrive.

Lastly, the development of the metaverse culture must be based on a solid security boundary, ordered rules and standards, and a correct ethical philosophy to ensure the bottom-line and orderly, sustainable operation of the metaverse. Bottom-line security includes the security of individual privacy, the data security of institutions/organizations, network security, etc., as well as whether the digital content itself violates the rules, constitution, protection of digital asset security, or has other issues. Through proper compliance with rules and standards, the metaverse can develop in an orderly manner, and an ethical philosophy that is consistent with the development of human civilization not only defines the ethical boundaries for the development of metaverse technology, but is also a prerequisite for the sustainable development of the metaverse.

Creation

| | |
|---|----|
| Three Forces Integrating Virtuality with Reality that Create the Metaverse | 13 |
| Creativity: Accelerating the creation of the virtual world | 13 |
| Connectivity: Realizing the connection between digital and physical worlds | 13 |
| Integration: Pushing the integration of virtuality and reality and intelligent development of the metaverse | 14 |
| The Three Infrastructures that Consolidate Metaverse Productivity | 14 |
| Engine: A low-threshold and cross-terminal creation environment | 14 |
| Algorithm: Accelerating "creation," achieving "connection" and promoting "integration" | 15 |
| Computing power: Supporting the massive computation of the metaverse | 15 |

Three Forces Integrating Virtuality with Reality that Create the Metaverse

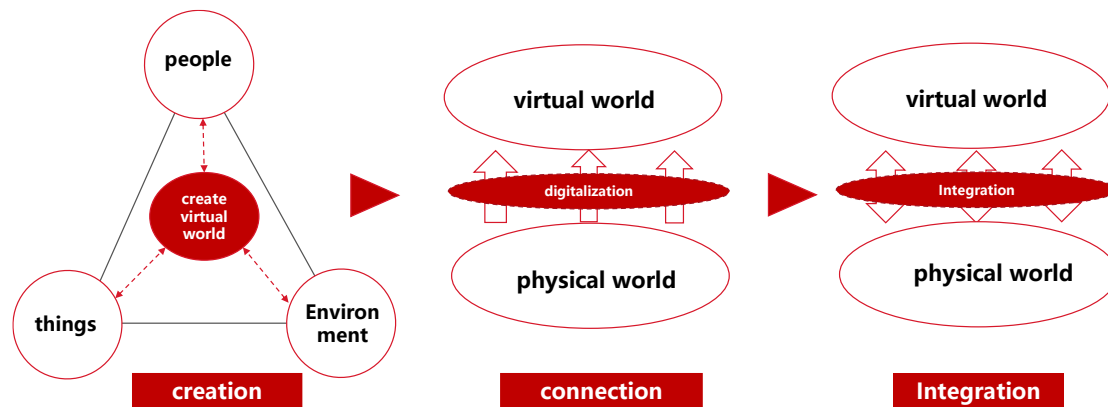


Figure: Three Forces Integrating Virtuality with Reality that Create the Metaverse

Creativity: Accelerating the creation of the virtual world

Based on 3D engines, we can conduct 3D digital modeling using the three fundamental elements of "people, things, and environments" to make the morphological appearance or animation effects of models more closely resemble the visual appearances of real scenes through simulation engine technologies like image rendering and physics, and then create VR/AR and other terminal experiences for users in combination with XR interaction engines. In short, we create a virtual world using 3D engines that runs parallel to the real world.

To further improve creation in the virtual world, 3D engines can also introduce AIGC (AI Generated Content) to quickly generate virtual content, lower the threshold for content production, and reduce the cost and cycle investment of content creation. This way, metaverse applications will no longer be limited to the gaming industry, and can be promoted by and applied to more industries and sectors on a broader scale. For instance, through the use of AIGC, NVIDIA's Canvas is able to quickly "imagine" and generate highly authentic scenes from only a few strokes of doodling or the input of text or voices. Merely from shooting photos of an individual, SenseTime can generate a highly-precise figure model in about a week, while traditional CG production requires at least several months and millions in costs.

Connectivity: Realizing the connection between digital and physical worlds

The virtual world we create via 3D engines, AIGC, or through other means has no direct connection with the real world, so we need to "relocate" real-world information to the virtual world so that the virtual world reflects how the real world works in a synchronous manner. This is precisely what AI is busy doing today—the digitalization of the real world, which is to say the original unstructured data collected from different IoT devices in the real world, in combination with AI IntelliSense, is transformed into structuralized data that machines can understand. Then the processed

data is projected into the virtual world to break the "dimensional wall," thus establishing a connection between the digital and physical worlds.

Integration: Pushing the integration of virtuality and reality and intelligent development of the metaverse

Based on the connection between virtuality and reality, and in combination with the real-time data projected by the physical world, the optimal strategies and decisions are used to create reverse instructions for optimization and operation in the physical world through low-cost trial-and-error simulations, tests, or other activities in the virtual world. This is the process by which the metaverse empowers economic development in the real world. For example, by connecting virtuality with reality, we can monitor road traffic on a "visualized" basis, and then, using AI decision-making algorithms, analyze and estimate real-time traffic data to simultaneously formulate the optimal traffic strategy, and finally optimize real-world traffic synchronously by controlling the traffic lights and other signals to ease traffic jams.

AI technology, which plays a significant role in the creation and development of the metaverse, accelerates the creation of the virtual world, connects the digital and physical worlds, and promotes the integration of virtuality with reality and the intelligent development of the metaverse.

The Three Infrastructures that Consolidate Metaverse Productivity

Engine: A low-threshold and cross-terminal creation environment

For an open metaverse, creation rights are certain to fall to users. An intelligent 3D engine with low thresholds and an open-source environment for creation and the integrating of AIGC functions will be the infrastructure for content creation in the metaverse, and it will be the basis for a large-scale creator/developer ecosystem. This engine will be what allows rapid response to new and massive requirements for the development of metaverse content. Meanwhile, a wider range of users can partake in created content because the engine allows content to be developed once and ported to a wide variety of terminals without further adaptation. For example, content created by a developer on Unity may be created once and deployed to over 20 interactive terminal types, including Windows, Mac, iOS, Android, PlayStation, Xbox, Nintendo Switch, as well as AR & VR platforms. SenseTime's SenseMARS Mixed Reality Platform is not only compatible with different forms of applications such as apps, applets, and H5, it also supports over 200 smartphones, tablets, AR/VR glasses, smart TVs, drones, and other terminal devices. Thanks to cross-terminal compatibility, creator/developer workload is also reduced, and users can access and experience metaverse content using any of end systems.

Algorithm: Accelerating "creation," achieving "connection" and promoting "integration"

As mentioned above, massive AI technology support is required for the creation, connection, and integration of the metaverse, and in turn to accelerate the "creation" of content and promote connection and integration between virtuality and reality. For example, a number of perception algorithms are involved in the connection between virtuality and reality, i.e., the digitalization of the real world. According to our calculations, a massive number of algorithm models is required for the digitalization of the entirety of the real world. Such algorithmic requirements must be matched with an efficient mode of algorithm production.

Therefore, we need to build a platform for algorithm production at the industry level. On one hand, the whole process, from data storage, annotation, training, and inference to deployment will be streamlined and standardized, thus shortening the cycle of algorithm innovation, improving the efficiency of algorithm production, and rapidly responding to the requirements for the digitalization of metaverse scenarios. Additionally, as regards the digitalization of multitudinous instances of low-frequency, long-tail "fragmented" scenarios, the generalization capacity afforded by algorithm infrastructure built around an "foundation model" effectively alleviates the problem of repeated modeling in fragmented development, and meets the requirements for digitalization of long-tail scenarios during the metaverse's construction, while at the same time reducing the development threshold.

Computing power: Supporting the massive computation of the metaverse

Computing power is the cornerstone of metaverse development. Massive computing resources are required to create a sense of reality, timeliness, intelligence, and content creation for the metaverse experience of the future. According to IDC estimates, by 2030, the total computing power required for the metaverse (including AI, VR/AR, IoT, blockchain, etc.) will be hundreds of times greater than the current scale. One of Intel's senior vice presidents said that computing power must experience a 1,000x increase if we want to realize a metaverse experience as shown in Avalanche or Ready Player One.

The era of massive computing power has arrived! Meanwhile, the deployment of computing will undergo structural reforms. First, traditional CPU-oriented computing architecture will not be able to meet the demands of real-time processing and the analysis of massive unstructured data generated by the consumption metaverse and industrial metaverse. Second, with the increasing demands for computing power, and due to restrictions on the development of network technology and the cost of network bandwidth, the deployment of edge computing will inevitably offset this deficiency. Therefore, upgrading metaverse computing will primarily be focused on the "cloud-edge-terminal" coordination model, which relies on intelligent (heterogeneous) computing (AIDC).

Empowerment

| | |
|--|----|
| SenseMARS Mixed Reality Platform: Engine Infrastructure that Creates and Designs the Metaverse | 17 |
| SenseMARS Avatar: "We/us" in the metaverse | 18 |
| SenseMARS Agent: "They/them" in the metaverse | 19 |
| SenseMARS Reconstruction: Digital reconstruction of the physical world | 19 |
| SenseCore Universal AI Platform: Supporting the algorithms and computing power infrastructure of the metaverse | 20 |

Through the SenseMARS Mixed Reality Platform, which is focused on "people, things, and environment" and uses AI technology, SenseTime empowers developers, with a low threshold, to create a metaverse world that integrates virtuality and reality efficiently, and to create immersive experiences with enhanced interaction and mixed reality. Additionally, SenseTime's SenseCore Universal AI Platform provides efficient algorithm and computing resources for creating and designing the metaverse to speed up its creation, strengthen the connection between virtuality and reality, and promote the integration and intelligent development of the virtual and real worlds.

SenseMARS Mixed Reality Platform: Engine Infrastructure that Creates and Designs the Metaverse



Figure: SenseMARS, engine infrastructure that creates and designs the metaverse

With our focus on "people, things, and environment", SenseMARS exports functions and services — SenseMARS Avatar, which rapidly generates virtual avatars to help people enter the metaverse and traverse the virtual world; SenseMARS Agent, which supports the development of digital humans and other smart agents that provide us with various smart services in the metaverse while also interacting with people in an intelligent manner; and SenseMARS Reconstruction, which achieves 3D digital reconstruction of the physical world and creates a virtual copy of the physical world.

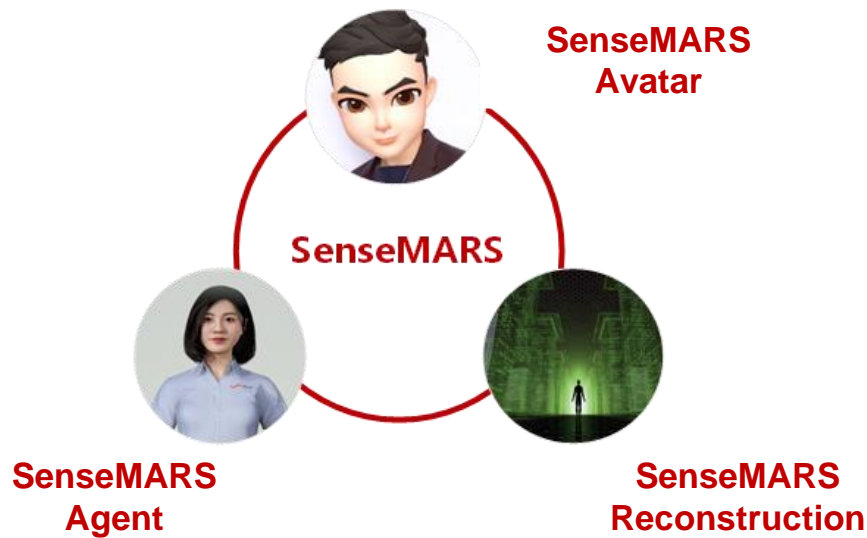


Figure: SenseMARS exports functions and services

SenseMARS Avatar: "We/us" in the metaverse

An avatar is the digital ID we use to enter the metaverse, and also our "second life" in the metaverse. In every different virtual scenario, we can choose virtual avatars of different styles and appearances to fully express our different personalities in the metaverse. SenseMARS Avatar is a critical engine that lets us efficiently create virtual avatars.

Based on SenseMARS Avatar, we can use personal images and AIGC to rapidly generate avatars of different styles such as anime, cartoons, simulated humans, and hyper-realistic 3D. Using the world's leading computer vision technology and AI motion analysis, we are also able to use ordinary RGB cameras to achieve motion capture without professional optical cameras or wearable sensors so that everyone can easily convert their body, face, movements, and language to their digital avatars in any metaverse.



Figure: avatars of different styles

SenseMARS Agent: "They/them" in the metaverse

Apart from "us" in the metaverse, there is also a group of intelligent "them," the intelligent virtual agents. "They" can interact with us in a smart manner and provide various smart services. SenseMARS Agent is the key technology engine for creating smart agents.

By integrating a series of AI technologies, including Computer Vision (CV), Automatic Speech Recognition (ASR), Natural Language Processing (NLP), Speech to Animation (STA), intelligent decision-making, and deep learning, we can equip a virtual smart agent in the metaverse with a "smart brain" that allows us to interact with "them". Digital humans, for example, can "understand" human language and communicate with us via language, facial expressions, and body movements. Plus, through training and learning the knowledge of different sectors, digital humans can become omniscient and act as our "super assistants" in every sector.



Figure: hyper-realistic 3D digital human

SenseMARS Reconstruction: Digital reconstruction of the physical world

We can create massive virtual scenarios in the metaverse with 3D reconstruction of the real world. SenseMARS Reconstruction is the key engine that allows us to rapidly duplicate the real world.

By relying on the integration of multiple algorithms (e.g., 3D semantic segmentation, MVS, etc.) and empowering developers to use consumer-grade mobile devices (cellphones, action cameras, drones, etc.), SenseMARS Reconstruction can efficiently reconstruct 3D models of the physical world and precisely duplicate them at the centimeter level. It can do this for everything from small objects to shopping malls, transportation hubs, and even cities. Furthermore, in combination with the precise

space mapping and visual localization of SenseMARS, visual content can be superimposed on the physical world via AR glasses, smart phones, and other terminals to realize accurate superimposition and seamless integration of the physical and virtual worlds.



Figure: 3D reconstruction of the real world

SenseCore Universal AI Platform: Supporting the algorithms and computing power infrastructure of the metaverse



Figure: SenseCore Universal AI Platform

The SenseCore Universal AI Platform provides underlying support to the AI models and computing resources required for the creation and design of the metaverse.

First, through our hyper-scale and intensified computing power deployment, SenseCore can reduce the costs associated with computing and AI model research & development (R&D).

As of June 30th, SenseTime has set up 23 intelligent computing centers in major regional markets that altogether offers computing power of 1170 Petaflops. Additionally, with a peak computing power of 3740 Petaflops, the Artificial Intelligence Data Center (AIDC) established in the Lingang Special Area of China

(Shanghai) Pilot Free Trade Zone at the start of 2022 is now one of the largest AI intelligent computing centers in all of Asia, and is able to meet the computing requirements of four hyper-scale smart cities, each with a population of 200 million. With our complete proprietary AI scheduling system and distributed AI storage system, SenseTime has created a simple, efficient, and uniform framework for AI application and development that increases the labor efficiency of AI development by 60 times, decreases the TCO (total cost of ownership) by over 70%, and accelerates the integration of AI and metaverse applications.

Second, SenseCore has connected all stages, including data processing, model training, high-performance inference & computing, and model deployment to achieve the mass production of AI models through standardized and automated processes.

In comparison with the industry's standard R&D cycle of several weeks, SenseCore empowers AI models for the entire production process that can greatly improve production efficiency and shorten R&D cycles to just several hours. Thanks to continuous refinements in productivity, the SenseCore AI studio is able to make training R&D 12 times more efficient and deployment R&D 40 times more efficient. It also supports multiple cloud and edge inference devices with algorithm models for inference optimization that increase inference performance by a factor of 10, making AI production R&D faster and easier. To date, SenseTime has produced over 49,000 AI models that cover the digital scenarios of multiple industry verticals.

Third, by supporting training and R&D for high-performance and precision foundation models, SenseCore can further accelerate the production of AI models to solve complex long-tail problems.

Based on SenseCore Universal AI Platform, SenseTime continues to invest in foundation model R&D, i.e., the production of generalized and universal pre-trained models through "massive computing power + big data" training, which further improves the production efficiency of AI models and provides digitalization solutions for massive long-tail scenarios. With close to 30 billion parameters to date, computer vision (CV) foundation models trained on SenseCore platform possess the largest number of parameters disclosed in the CV sector.

Cases

| | |
|--|----|
| Case 1: Asia's "unmanned" help desk created by AEON | 23 |
| Case 2: Zepeto - Custom avatars through rapid face molding | 25 |
| Case 3: AR Navigation at Suvarnabhumi Airport, Thailand | 25 |
| Case 4: Watching games via AR interaction in a Japanese baseball stadium | 27 |
| Case 5: Riyadh Season's immersive AR journey | 28 |
| Case 6: Schwarzkopf - AR hair-dyeing trial | 31 |
| Case 7: AR Digital cultural and creative platform | 32 |

Based on the SenseMARS engine infrastructure, as well as the algorithm and computing power provided by SenseCore Universal AI Platform, SenseTime has created a foundation for the metaverse, and empowers all sectors and industries in their efforts to effectively create and design in the metaverse as they integrate virtuality and reality. SenseTime is also actively exploring innovative metaverse scenarios in the Asia-Pacific region and worldwide to empower the sustainable development of local digital economies and deliver a more immersive experience through virtuality-reality integration.

Case 1: Asia's "unmanned" help desk created by AEON



Figure: the digital assistant "Xiaotang"

AEON is a leading general retail and service group in Asia with headquarters in Japan, China, and Southeast Asia. AEON is primarily engaged in shopping center and full-scale retail industry operations (shopping malls, food supermarkets), while also managing additional business such as specialty stores, financial services, property services, and convenience stores.

In the presence of new retail trends, traditional businesses like AEON are actively seeking their own digital transformation, and are using new technology and new philosophies to reconstruct the relationships between "people, things, and environment" in such a way that both commercial flows and commercial scenarios can expand. Starting with services and experiences for "people", there were three pain points to address in the business operations of AEON stores:

1) The stores themselves encompass a large floor area, are numerous, and are widely distributed. Multimedia screens are located on each floor to provide map navigation, serve as billboards, facilitate event marketing, and provide other information. There are also physical billboards in nearly every corner of the retail spaces. However, the videos playing on screens and billboard displays offered little in terms of customer interaction and resulted in a subpar service experience. Additionally, it was difficult to gain precise information on marketing conversion rate.

2) As the major platform for offline service at stores, the help desk faced difficulties such as high labor costs, high staff turnover, inconsistent service quality, and low extension in services.

3) Due to new retail trends, the stores urgently needed a more direct connection with customers to more accurately grasp their consumption needs and interests, as well as a more scientific method of digitalization to empower their operations-related decision making and improve their services.

Based on the problems described above and the technical capabilities of the SenseMARS Mixed Reality Platform, AEON Mall Guangzhou Xintang and SenseTime jointly created the digital assistant Xiaotang, which provided customer consultation, navigation, shopping guidance, and other services.

As a virtual customer service provider, Xiaotang could provide an accurate and rapid overview of locations within stores, navigation of parking routes, credit exchange, and other consultation services. When problems arose that "overloaded" the program, background managers could remotely take over in a timely manner. It is notable that 80% of such problems had to do with finding lost persons or items where a human to verify the relationship or ownership.

As a virtual shopping guide, Xiaotang was able to determine the actual requirements of customers, and more effectively and accurately recommend new products, and communicate new store openings, special offers, and other information related to shopping malls or stores, which helped improve marketing conversion.

At present, Xiaotang's knowledge base covers greetings, the entertainment and leisure preferences of users, product sales information, while also providing store navigation and other similar content. Every day, it answers over 1,170 questions, serves over 100 customers, and answers questions of over 10 rounds. It comprehends 92.7% of customers' daily questions, it has an accuracy rate of 95%, and it gradually improves itself from daily "learning".

AEON is now establishing this sort of "unmanned" help desk all over Asia, which, according to early estimates, can help shopping malls save over 50% in labor costs. As more "Xiaotangs" replace traditional service windows and become an important point of offline connection with users, user data is comprehensively collected by multiple facilities within the shopping malls to create a form of multi-point or multi-dimension interaction, thus establishing a private customer traffic and information database for the shopping malls. This provides them with the basis for scientific data analysis to further refine their services and more efficiently manage store brands in the mall.

Case 2: Zepeto - Custom avatars through rapid face molding



Figure: Zepeto uses AI to generate a personal avatar

As the largest metaverse platform in Asia, Zepeto is a global phenomenon. It has accumulated over 300 million users since the social application was released online by the Korean company Snow in 2018. Users can "mold" their personal animation image and decorate their personal space and create their own virtual avatars to display their individual interests and lifestyle and build relationships with strangers.

For an "image"-oriented entertainment social application, a lack of personalized characteristics and sameness are unacceptable. To meet user demands for virtual avatar creation, such as face molding and costumes, SenseTime and Snow have created the first face molding and costuming plan for the virtual world. By taking photos to identify facial characteristics and applying AI or augmented reality (AR) technology, users can rapidly create a virtual avatar that mirrors their own image and then use it as a basis to shape their unique avatar by adjusting the facial contours, eyes, nose, mouth, and ears.

Case 3: AR Navigation at Suvarnabhumi Airport, Thailand

For many years, the government of Thailand has had its focus on digital development. Since 2016, the government of Thailand has guided its digital economy under the "Thailand 4.0" scheme, which is a 20-year road map leading Thailand toward the goal of being a "value-oriented and innovation-driven economy". The scheme focuses on digital improvements that can enhance the lives, productivity, and efficiency of the Thai people.

"Thailand 4.0" stipulates that airports, as important infrastructure for air transportation and cities, play a key role in the development of the digital economy and the digital transformation of the nation. According to Mr. Nitinai Sirismatthakarn, the Executive Director of the Airports of Thailand (AOT), by promoting technological solutions,

AOT improves its services to make airports more vibrant and ultimately bring about the digital transition. "This is a major challenge for the Ministry of Transport".

Suvarnabhumi Airport is one of the six airports under the AOT. It covers an area of 32 km² and ranks first in passenger capacity and aircraft movements in Thailand. It is one of the most important aviation hubs in Southeast Asia, as well as in the Asia-Pacific region. In 2019, SenseTime and SKY ICT cooperated to make digital improvements to the airport's massive physical space of 500,000 m² by applying AI and mixed reality (MR) technology to optimize the service experience of passengers at the airport.

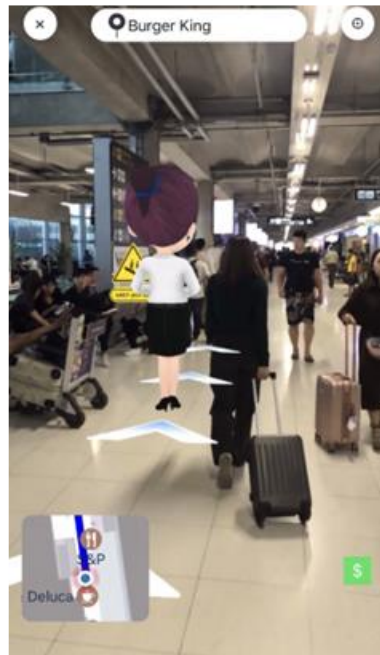


Figure: AR Navigation at the airport

AR Navigation: Using digital restructuring, visual locating, and MR technology, a passenger's real-time position at the airport can be ascertained, making AR navigation possible. In combination with mobile AOT, after reaching the airport, passengers can see their position and access a convenient route-guidance service by turning on AR navigation in the app and scanning their surroundings in the airport. By following the AR arrows and virtual guide, passengers can easily reach their destination to access visas, currency exchange, taxi rentals, shopping, and other airport services. Meanwhile, AR navigation is linked to the passenger boarding system, allowing navigation directly to boarding gates. Being more efficient than traditional information desks or 2D site maps, AR navigation provides passengers with accurate and uninterrupted service and helps to conserve their precious and limited time for other activities at the airport.

AR Marketing: The AR navigation includes AR billboards that display the stores along the passenger's route. Users can obtain real-time information about special offers, recommended goods, per capita consumption, and other key information about the stores. This AR marketing provides an intuitive and rapid reference for passengers shopping in the stores, makes their shopping experience more convenient, and furthers the marketing conversion of airport stores.



Figure: AR Billboards at the airport

Realizing precise navigation across large spaces at the airport using AI or MR technology is a major innovation for the aviation industry. Most traditional navigation services for airports are based on GPS positioning or Bluetooth beacons, which produce unavoidable problems like major positioning errors, high hardware costs, and complex maintenance. Our industry-leading CV and MR based on the SenseMARS platform can rapidly create airport navigation services that are useful in multiple, integrative service tools such as interior positioning in large spaces, navigation in complex exterior areas, emergency positioning, cross-floor positioning, and underground parking lot positioning, all of which facilitate the digital transformation and creation of smart airports.

Case 4: Watching games via AR interaction in a Japanese baseball stadium

In Japan, baseball is the most popular national sport, followed by sumo wrestling. According to data released by the professional baseball authority in Japan, the number of spectators watching NPB (Nippon Professional Baseball Organization) games at home stadiums was at one point over 26.53 million, 30,928 for every game. However, the sudden emergence of the COVID-19 epidemic cast a dark shadow over the baseball industry in Japan. With weakened sporting event consumption, shuttered institutions and stadiums, and postponed games, the industry faced unprecedented challenges. In 2020, the average number of spectators declined to 7,805, an annual decrease of 82%. The drastic decline in spectators watching baseball games led to reduced revenue for stadiums and baseball clubs.

In May 2022, SenseTime and its local partners in Japan provided AI and MR support to a well-known Japanese baseball stadium. By providing an AR platform and integrating interesting AR effects at the stadium, we attracted more people by delivering a unique set of interactions for spectators to make post-epidemic baseball games livelier and more immersive. Currently, the AR effects appear mostly at two parts of the stadium: the entrance and the field.

At the entrance to the stadium, a human-controlled AR baseball girl avatar serves as the game hostess. By scanning the scene at the entrance with H5, spectators can interact with the baseball girl avatar, take photos with her, and post them on social media to create additional discussion related to the game. Meanwhile, MR technology is used to display AR advertisements on the live-action background behind the baseball girl, which converts people's attention to business value.

On the AR displays at mobile phones at the venue, spectators can see AR balloons being "released", allowing them to experience the celebrations on the field without the environmental pollution of actually releasing balloons. They can also see AR-enhanced ball motion, velocity, and other pitch analysis data to further enhance the game experience.



Figure: spectators can see AR balloons being "released"

During games in May 2022, the AR application at this stadium was used by 2,000 people nearly 4,000 times over a period of just three days, which generated increased attention for the games through attendees sharing their experience on social media.

Case 5: Riyadh Season's immersive AR journey

As an important part of Vision 2030, Saudi Arabia has been promoting cultural tourism development in recent years. Themed cultural tourism seasons are held in key tourism cities to put the country's natural landscapes and cultural customs on display for tourists from all over the world. Riyadh Season is the biggest cultural and entertainment event in Saudi Arabia, perhaps even in the Middle East, and also plays an important role in the development of the digital cultural & tourism industry, and Vision 2030, in Saudi Arabia. With an event zone covering an area of 900,000 m² across 14 themed zones, and including about 7,500 activities, Riyadh Season is a festival of entertainment that integrates music, art, culture, catering, and more. According to data from the General Entertainment Authority (GEA), there are over 11 million Riyadh Season visitors, including 1.6 million from abroad.

To attract more tourists, to make Riyadh Season even more appealing and popular around the world, and to serve the sustainable development of Saudi Arabia's cultural tourism industry, SenseTime and Sela, a sports industry management company based

in Saudi Arabia, cooperated to empower Riyadh Season with AI and MR technology based on SenseMARS platform, creating a brilliant cultural and entertainment experience that was markedly more immersive and interactive for local and global visitors across five zones - Riyadh Boulevard, Combat Field, Winter Wonderland, Safari, and Riyadh Front. The project currently features four functional modules, including AR navigation, AR themed routes, AR spots, and AR marketing.

AR Navigation: Since the tourism park covers a large area across multiple zones, SenseTime uses AI and MR technology in the park to run a live AR navigation service covering tourist spots, prayer rooms, medical services, public toilets, and other services. Simply by turning on the mobile app, the AR navigation, which is available anytime, provides visitors with a more convenient and engaging experience by helping them find locations quickly and conveniently and access electronic guidance. Meanwhile, real-time analysis on the number of tourists in queues using visual perception technology and AR navigation allows tourists to check the queues and better arrange their schedules during their trip.

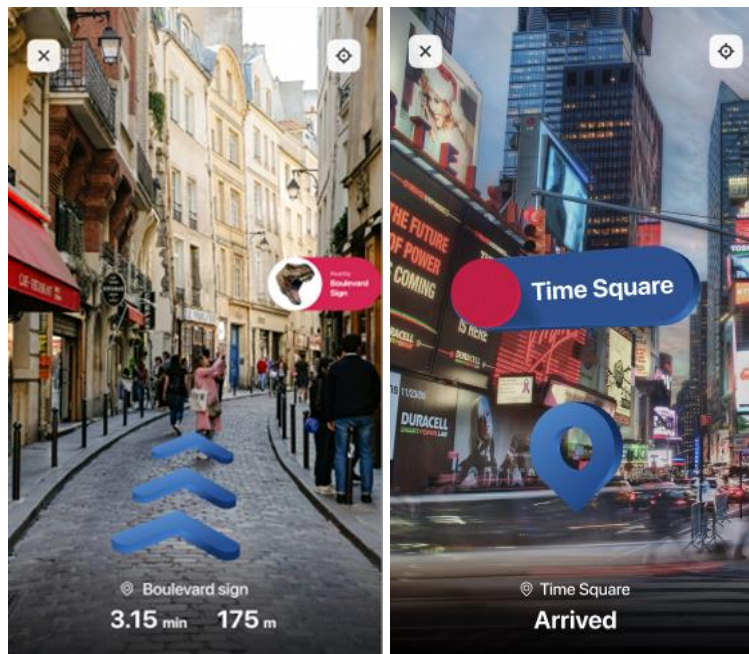


Figure: AR Navigation at Time Square area

AR Spots: Supported by SenseTime's MR technology, AR spots were established at two locations in Time Square - Dinosaurs and Fountains. By scanning the Dinosaurs location in the mobile app, tourists can see a virtual T-Rex striding back and forth, looking around, and roaring. By scanning the Fountains locations, tourists can see an ocean in the sky where a "whale" leaps high from a fountain and splashes, producing a stunning visual impact for tourists. Tourists can take photos at the AR spots and post them on social media to draw more attention to the park.



Figure: AR spots at Dinosaurs and Fountains areas

AR Marketing: While providing AR navigation for tourists, the park also includes AR billboards along the route to provide recommendations on the stores around the locations on the route, where store marketing can include special offers, discounted goods, and other recommendations. Combining AR marketing and traditional marketing improves the conversion efficiency of shopping.



Figure: AR billboards

AR Themed routes: Due to the complex layout of Riyadh Boulevard, the park provides AR themed routes based on AR navigation, which connect AR spots and AR marketing sites together into a suggested tourism route, making it easier for tourists to find and visit sites of interest or promoted locations and providing store marketing in the form of AR clock-in and puzzle games through which they can receive coupons.

Using the SenseMARS platform, the park provides global tourists with an efficient and convenient tour guide experience using a rational visual display, while the vivid,

engrossing, immersive, and interactive design gives tourists an opportunity to experience and embrace Saudi Arabia from a different and engaging perspective, thereby adding more appeal and content to cultural and tourism events.

Case 6: Schwarzkopf - AR hair-dyeing trial

The "gap between ideals and reality" presents an important problem for people who color their hair in pursuit of fashion. Is a color suitable for your skin tone? The color may look nice in theory, but why does the actual result look so different?

In cooperation with SenseTime, the 140-year-old German company Schwarzkopf provides its customers with an immersive AR hair-dyeing trial. Customers can try different hair colors, virtually, via a web application, mainstream third-party platforms, and other online channels or an offline AR make-up mirror. The direct and accurate display of complete dyeing results on the screen allows customers to be more confident about their color choice. The AR hair-dyeing trial also supports before and after comparison so that customers can clearly see the difference.

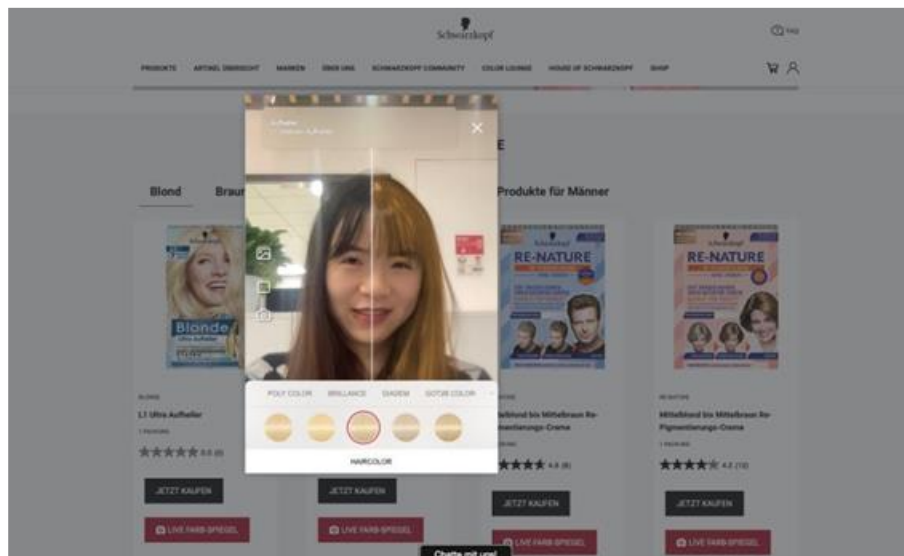


Figure: The AR hair-dyeing trial also supports before and after comparison

At the same time, using SenseTime's AI technology, Schwarzkopf's product design staff only need to enter RGB color values to automatically generate AR hair-dyeing results with realistic textures and sheen, after which no subsequent coloring adjustment is required. This makes Schwarzkopf's business operations more efficient and reduces the time to release new hair-dyeing products by 90%.

"Precise image recognition and seamless motion tracking and rendering are needed for AR hair-dyeing trials," said John Gao, global CTO of Henkel dx (a digital transformation business). Using AI + AR and other SenseTime technologies, Henkel can display the exact colors of hair dye for customers, and also reflect the multi-dimensional changes in the shine and texture of hair. The technology both resolves key pain points for consumers and helps them make quick purchase decisions, thus optimizing the shopping experience.

Case 7: AR Digital cultural and creative platform

Gen Z is the demographic that drives the cultural and creative market, and the individuals who make up Gen Z, colloquially known as Zoomers, impact models of consumption through their preference for interactive consumption over the one-way purchase and use of goods. However, traditional culture products, despite the certain commercial value they possess thanks to the IP held by cultural and creative institutions, regardless of whether they are released in a physical or digital form, are primarily products intended for either display or collected. The low level of interaction in these formats makes it difficult for traditional culture to build effective emotional connections with new consumer groups and convey cultural philosophies to younger audiences.

SenseTime and Dunhuang Cultural Creativity (a brand under the Dunhuang Culture and Tourism Group) collaborated to release "A Thousand Years in an Instant - A Limited Digital Mural of the Dunhuang Nine-Colored Deer", the first NFT collectible of its kind. The "hook" for this piece was its interactive AR digital creativity. The physical work of art was used as a vessel to combine and convey traditional culture, modern technology, and avant-garde interactivity in a format that appeals to Zoomer consumers with the goal of leaving a deeper impression of the magnificent cultural relics of Dunhuang.

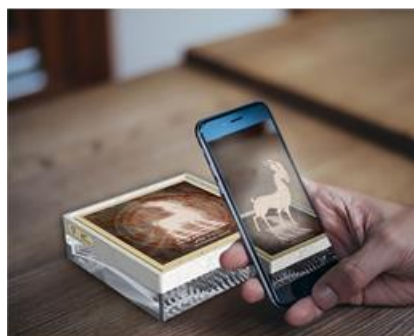


Figure: AR effect of "A Thousand Years in an Instant - A Limited Digital Mural of the Dunhuang Nine-Colored Deer"

The Nine-Colored Deer can be physically displayed, but the piece also integrates the innovative and immersive experience provided by SenseTime's AI + AR technology. By scanning the physical object with a related mobile app, a "portal" opens into the virtual recreation of the Mogao Grotto No. 257 at the Dunhuang Mogao Caves. Here, the Nine-Colored Deer mural emerges, and with gently flowing digital light effects, appears in all its former glory. The nine-colored deer in the mural appears to have been reborn, and the sounds of its movements along with its robust figure allow people to relive this classic fairy tale.

Epilogue

As a progressive process, metaverse development may be divided into four phases: the fragmented 1.0 Era, the 2.0 Era of systematization, the 3.0 Era of ecology creation, and the 4.0 Era of integration.

The 1.0 Era, the early stage of metaverse development, is mainly focused on "point-based" innovations in applications and content, in other words, fragmented scenarios. After gaining experience and refining their technology, enterprises will be able to realize sustained closed-loop value, at which point they would increase their investment in metaverse technology applications, and then, as they connect the dots, will go from "individual points" to a "line" as they link all their business scenarios and enter the 2.0 Era of systematization. As data standards are unified within the industry and data barriers are removed, there will be a vertical connection of elements between upstream and downstream enterprises within the industry, or between enterprises within the ecology. In the 3.0 Era of ecology creation, the boundaries of the metaverse will continue to expand. At present, our hope is to one day realize the 4.0 Era of integration, so that people can freely access and traverse the various regions of the metaverse whenever and wherever the need arises.

The journey from 1.0 to 4.0 is a long-term and progressive process. Technology plays an important role in this process, but other factors such as ethical governance, privacy and data security in the metaverse, the integration of real and virtual economic systems, and the organizational models of industry resources may also impact metaverse development. Currently, there is no clearly defined path to an advanced stage of development. Therefore, the administrations, enterprises, colleges, and diverse participants in Asia-Pacific countries need to feel their way through and jointly seek answers to questions as they proceed.

Report Advisory

| | |
|---------------|--|
| James Ong | Dean of the Artificial Intelligence International Institute Secretary General of Tech4SDG |
| Hiromi Komuro | Director General of the Japan International Metaverse Association |
| Fan Yang | Cofounder and Vice President, SenseTime |
| Jeff Shi | President of Asia Pacific Business Group, SenseTime |

Author



Tian Feng

Secretary General of Tech4SDG
Dean, SenseTime Intelligent Industry
Research Institute
E-mail: tianfeng@sensetime.com



Yang Yan

Director of Strategy and Ecology
Research, SenseTime Intelligent
Industry Research Institute
SenseTime
E-mail: yangyan1@sensetime.com