



2023

**WHITE PAPER OF
TELECOMS AI INDUSTRY
DEVELOPMENT**

2023.10

Acknowledgement

This White Paper has been produced as a collective effort within the SCFA AI WG, and on its behalf the following editing team (**listed in alphabetical order**):

China Mobile:

Di Jin, Jinglei Liu, Kaixi Liu, Lei Lei, Lin Zhu, Lingli Deng, Na Wu, Qihui Zhao, Sheng Wang, Yanpeng Sun, Yanping Liang, Yan Zhu, Yanjiao Wang, Yichun Sun, Zhijie Ren

KT:

Gyuseob Lee, Heeyoon Hong, Jaehyoung Park, Keehoon Sung, Kyunghyuk Yoon, Soonmin Bae, Soyeon Kim

NTT DOCOMO:

Eri Koide, Harunobu Maejima, Hiroki Ishizuka, Hisato Sugiyama, Izumi Haruta, Katsuhiko Ogikubo, Kotaro Abe, Kuan-yin Liu, Kohei Kawase, Midori Onogi, Noboru Miyamoto, Rina Sato, Takafumi Yamauchi, Takefumi Naganuma, Takeshi Kato, Takuya Kitade, Tatsuya Nakamura, Tomohiro Mimura, Shin Ishiguro, Shoichi Horiguchi, Satsuki Kumatani

SCFA AI WG

China Mobile Contact: liukaixi@chinamobile.com

KT Contact: gyuseob.lee@kt.com

NTT Docomo Contact: kuanyin.liu.bt@nttdocomo.com

Contents

1. Executive Summary	2
2. Overview of Development of Telecom AI Industry	3
2.1 Overview of global AI industry	3
2.2 Outlook of Telecom AI	4
3. Exploration and Progress of AI in Telecom Industry	5
3.1 Leading CSPs accelerate AI deploying, not only on network	5
3.2 Key Enabling Technologies	17
3.3 Summary	21
4. Best Practices of CSP	22
4.1 China Mobile Practices and Use Cases	22
4.2 NTT DOCOMO Practices and Use Cases	28
4.3 KT Practices and Use Cases	36
5. Industry Development Initiatives	42
6. Abbreviations and Acronyms	43

1. Executive Summary

This white paper is a joint release of China Mobile, NTT DOCOMO and KT. By revealing the development pattern and progress of global telecom AI industry, it aims to provide a CSP's perspective on the thinking and exploration along the transformation journey, on how telcos reshaping their roles under the rapid development and breakthroughs brought by AI. It further introduces the innovative use cases and best practices of telco in AI domain and intelligent network, showing the latest integration and development progress of AI in the telecom industry after AI technology entering large-scale deployment and monetization stage, and the profound changes it has brought to the whole industry.

With continuous evolution of telecom AI, we believe that CSPs will further become model supplier, model hub host and intelligence service operator. In the near future, telecom operators will leverage industry-specific advantages to continuously enhance collaboration with partners from the upstream and downstream of industry chain, as well as other different industries, creating a open AI ecosystem and a robust value chain fitting for the digital economy.

2. Overview of Development of Telecom AI Industry

2.1 Overview of global AI industry

Since December 2022, ChatGPT, launched by OpenAI, has presented a giant leap in understanding and generation ability, attracting worldwide attention and starting an AI gold rush. The Transformer-based LLM behind it presents a new way of energy and information integration, a new paradigm of industrial development, as well as a new round of scientific and technological revolution led by General Artificial Intelligence. Just as Newton's three laws are of significance to physics, the LLMs also brings the era of grand unification in AI.

Large Models will become a new type of infrastructure, with their role as the base of economic and social development becoming increasingly prominent. Great potential of LLM has been proved in search, work, intelligent interaction, AIGC, production workflow evolution, cost reduction and efficiency improvement of traditional industries. With the development of AGI represented by LLM, it is believed that AI services in the future will be as convenient, universal, safe and low-cost as today's water, electricity and communication services, deeply change people's lives and the society. In the recent years, there has been an increasing demand for computing force and data, as the number of intelligent computing centers and data centers steadily grows. According to the latest forecast¹, the revenue of AI semiconductors will increase by 25.6% and reach \$67.1 billion in 2024, maintaining an over 10% growth. By 2027, the revenue of AI chips is expected to be more than twice the market size in 2023, reaching \$119.4 billion.

AI will give birth to new business models, with digital intelligent transformation of traditional industries ushering in a new turning point. Without data being fully exploited, big data can only mean big cost and may not be of great value. While the emerging AGI makes the full mining of data value become possible on technique's standing. Industry private data, as the underlying support for AI model training, will enable the profound integration of large foundation model with industry-specific technology stack, form new barrier through Data Flywheel², realize fundamental transformation from X+AI to AI+X. Industry large models will be developed with fully utilization of domain-specific knowledge, expert experiences and production

¹ Gartner. 《Forecast:AI Semiconductors, Worldwide,2021-2027,2Q23 Update》 2023.08

² Data Flywheel:Refers to the process of continuously collecting, analyzing and utilizing data to realize the amplification and circular promotion of the data value. The data flywheel includes four stages, namely Acquisition, Activation, Application and Intelligence.

data. Large Models will further become the intelligent core of all industries and domains, to reshape the workflow, process, user experiences and business value, therefore transforming the intelligence into actual productivity, accelerating the monetization.

Large models are bringing changes of industrial development paradigm, forming a new ecology, promoting the division of labor in AI industry chain to become more clear. The digital intelligent transformation of traditional enterprises will also usher in new opportunities.

2.2 Outlook of Telecoms AI

The second curve of leading CSPs witness a rapid growth since digital transformation begins, with emerging services accounts for increasing portion in revenue. The telecom industry is going through a full upgrading on the new track of digital economy, AI and more fields, and has grown into a new positioning of innovative tech companies. The market has been reaching on a consensus that CSPs are becoming a major force of digital economy. According to Valuates, the global telecom AI market will grow from \$1.2 billion in 2020 to \$15 billion in 2027, with a compound annual growth of 42.6%.

In the upcoming AGI era, building a solid AI infrastructure and creating universal and safe AI services and operation abilities is the basis for achieving inclusive, ubiquitous intelligence, as well as enabling highly efficient and large-scale industry deployment. Telcos need a new paradigm towards intelligent operation, which can be summarized into 3 new patterns.

- 1. A new operation and supply pattern, enabling the operation and supply of broad scenario full-range elements.** Taking computing power, network, data, models, abilities and other intelligent elements as the operation and supply objects, integrated management of various elements can be realized through holistic planning, deployment, maintenance and optimization, etc. Intelligent service capabilities can be developed through integrated orchestration, scheduling and built-in security. Ubiquitous intelligent service can be provided for individual users and all walks of life through a systematic operation of all elements.
- 2. A new business pattern, accelerating transformation from X+AI to AI+X.** The system architecture, workflow and user experience of services today will be redesigned with the large model as the core. Not only the existing services and business will be re-constructed, but brand-new business forms will be surely created.

3. A new industrial cooperation pattern, with collaborative innovation as the goal. In AGI era, all industries will take the large foundation model as the base, combined with domain-specific knowledge, expert experience and production data , to develop industry-standard large models by collaboration. With further testing and improvement in the actual production environment, so as to truly achieve productivity leap through AI.

3. Exploration and Progress of AI in Telecom Industry

3.1 Leading CSPs accelerate AI deployment, not only on network

CSPs have rich scenarios for large-scale application of AI technology, massive and diverse high-quality data, large-scale cloud-based infrastructure, and user resources necessary for AI deployment and monetization, so they have inherited advantages in investing in AI R&D. Leading operators have long been aware of the disruptive potential of AI and started a comprehensive layout. Their exploration has been achieving and developed a differential competence along the digital intelligent transformation journey.

3.1.1 China Mobile: Open, Efficient, Secure and controllable intelligent service supply.

China Mobile aims to develop " Grand AI Platform" towards AGI realization, which is positioned as a new AI infrastructure, as well as a key carrier to realize the innovation of intelligent service operation paradigm. It includes large-scale intelligent computing center, data aggregation platform, AI model training and inference service platform, large foundation model towards national and social economic entities, industrial large models, aims to provide open, efficient, secure and controllable intelligent services based on ubiquitous computing force network.

So far, China Mobile has been achieving in the following aspects:

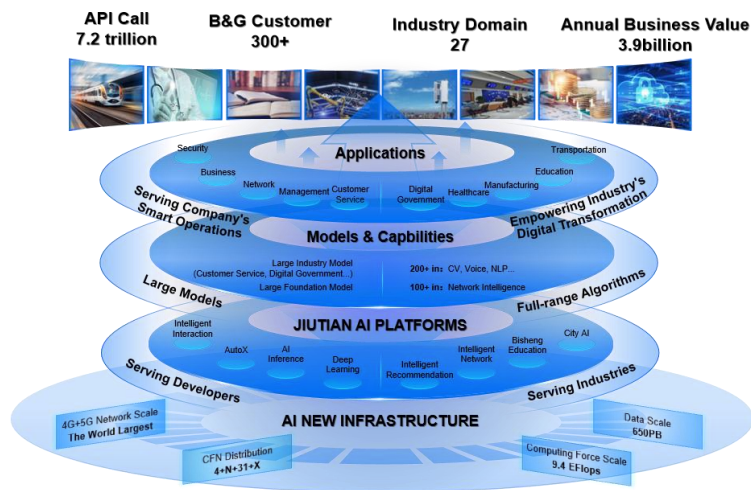


Figure 1 CMCC JIUTIAN AI Products

1) Solid network, computing force and data infrastructure.

- World's largest communication network with widest coverage and most user, with over 6 million 4/5G base stations and 370 million users covered by Gigabit broadband.
- Advanced computing force network, integrating cloud/edge computing, network and IT /CT domain into one important infrastructure, forming a "4+N+31+X" data center layout.
- Delay between computing nodes is shortened to 20ms across the nation, 5ms across the province and 1ms across the city. More than 650PB high-value data are gathered. A distributed collaborative computing platform is built.

2) “XINHE” Computing Native Platform, Enable Cross-Architecture Deployment and Migration of AI Applications

In 2022, collaborating with the industry, China Mobile originally proposed computing native technical system – Computing Native Application Migration (CNAM), which contains four key technologies including unified programming model, cross architecture compiling, standard intermediate language, and adaptive runtime. CNAM’s goal is to shield the difference of underlying heterogeneous hardware, break the tightly locked ecosystem of “AI framework + AI toolchain + AI hardware”, reduce the difficulty of cross-architecture deployment and migration of AI applications, and create an integrated intelligent computing ecosystem. Based on computing native technical system, China Mobile launched “XINHE” plan in June 2023 to develop “XINHE” computing native platform together with partners in industry, universities, and research institutes. “XINHE” platform 1.0 has been released during China Mobile Global Partners Conference in October 2023. Four core components of this platform are unified programming environment, CUDA to SYCL code porting toolkit, cross

architecture compiler, and adaptive runtime. It provides developers with end-to-end consistent development and deployment experience over heterogeneous computing resources. Now, “XINHE” platform has enabled typical AI applications, such as image processing and video analysis, to be deployed and migrating among multi-vendor and multi-arch AI processors. With “XINHE” platform, China Mobile opened up a new software-driven path to promote intelligent computing ecology, achieved the vision of “code once, build once, run anywhere”.

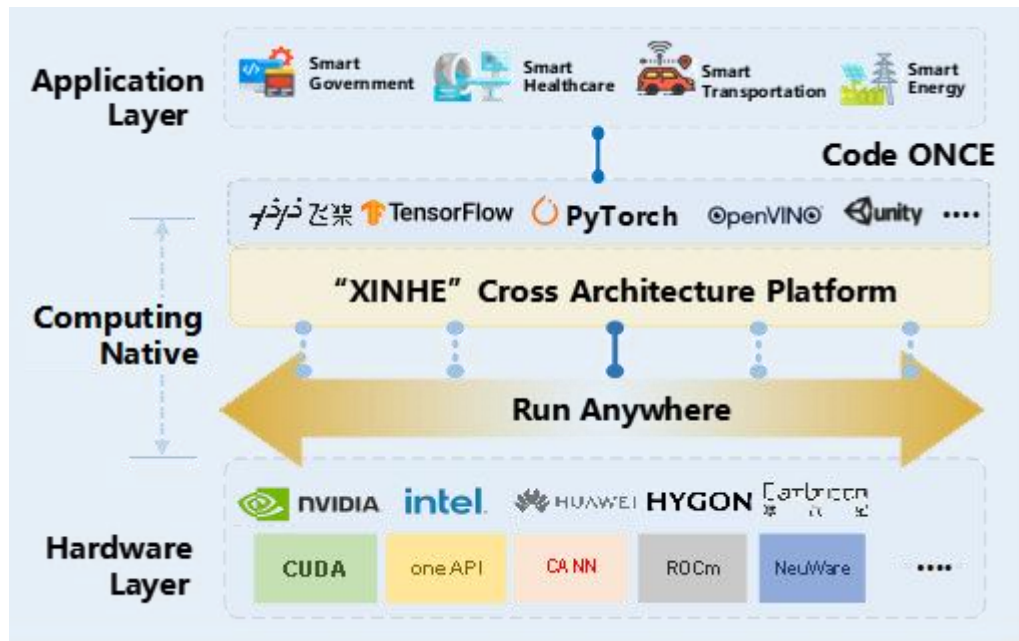


Figure 2 “XINHE” Computing Native Cross Architecture Platform

3) "JIUTIAN" AI Platform centered new intelligent computing engine.

The platform supports diversified intelligent computing and deep learning framework, cross-domain heterogeneous computing based efficient scheduling, high-performance model training acceleration with thousands of cards, and cloud-edge integrated training. The platform can now provide large-scale model training and inference and enable new MaaS. China Mobile has formed a comprehensive industrial-level intelligent service capability from PaaS, AaaS to large-scale deployment.

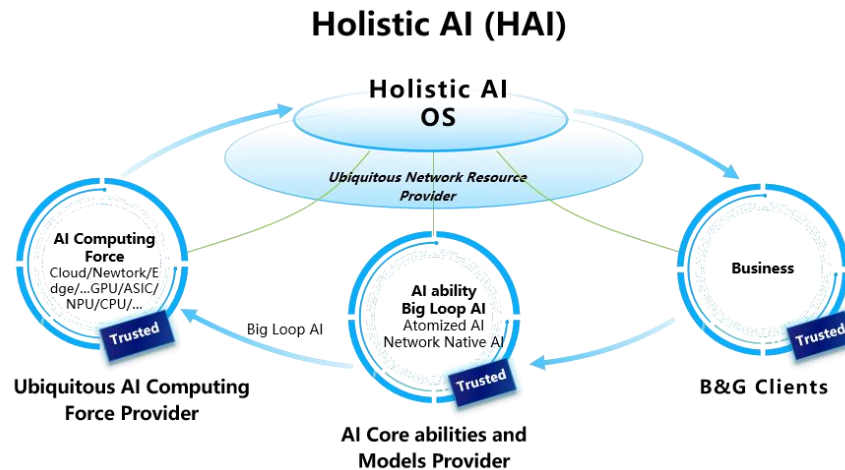


Figure 3 CMCC Holistic AI

4) "Holistic AI" theory innovation.

Holistic AI is originally put forward by CMCC to solve the challenges of high deployment cost, low efficiency and difficult localization in development of large model and AI. As the technical core to realize intelligent services operation, it mainly studies the theory, technology, mechanism, paradigm and framework needed for systematic reconstruction of AI technology. The goal is to flexibly and efficiently schedule, configure and monitor AI abilities according to the service requirements in an open environment, relying on ubiquitous communication networks and intelligent computing power. So that intelligent services can run and be supplied with most reasonable resources, as well as with trustworthiness, control and security.

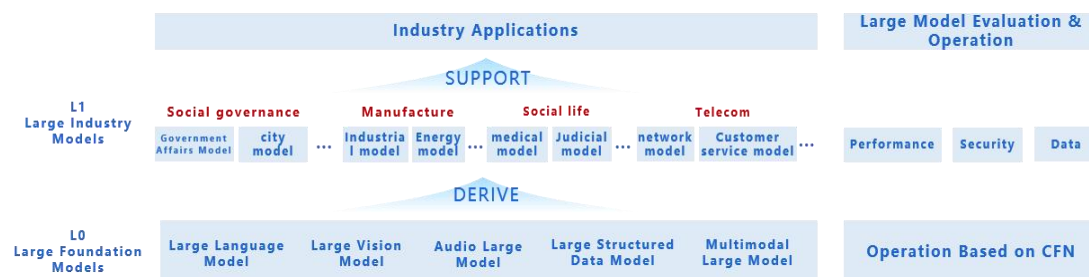


Figure 4 CMCC JIUTIAN Large Model System

5) JIUTIAN large model system supports a trinity pattern as a model supplier, model hub host and intelligent service operator.

China Mobile JIUTIAN large model system aims at serving the high-quality development of the national economy, fully integrates and stimulates the value of computing force, knowledge and data, and creates a new enabling paradigm of large foundation model as the L0 base, industry large model as the L1 development direction. As shown in Figure 3, the large foundation models ensure the consistency with mainstream values through technology innovations such as model design and

training, and has flexible and detachable structure and performance oriented to different industries, supporting flexible and low-cost deployment in various environments. The industry large models use the proprietary data of the industry for training, which can be integrated into the major production process and embedded into the complex production system, conforms to the industry-specific indicators, and has the adaptive ability and continuous industry knowledge learning ability in a highly dynamic environment. In addition, based on the collaborative empowerment mechanism of L0+L1, China Mobile has initially built JIUTIAN large model evaluation system, which has the characteristics of multi-level, multi-dimensional, multi-task, multi-mode and multi-index, to evaluate the performance of large models scientifically, fairly and objectively. The purpose is to know and understand all large models and promote the iterative optimization of self-developed ones, therefore providing a strong impetus for the development of the AI industry. Thirdly, China Mobile is committed to be the intelligent service operator by providing secure, efficient, flexible, universal and low marginal cost intelligent services, through unified scheduling of computing power, network and AI abilities and efficient service optimization and resource matching.

6) Joint efforts in new AI ecology development.

China Mobile has started a collaboration project to introduce the outstanding AI technologies and solutions, AI software and hardware, and AI integration abilities. Relying on the "JIUTIAN" AI product system and the advantages in computing, data, scenarios and capital, a new ecology blueprint is jointly made to accelerate the large-scale innovation and deployment. So far, 58 industry partners in China have joined the program for collaboration on R&D, products, marketing and standardization, including :10 universities and institutions such as THU and XJTU, 30 AI software and hardware enterprises,11 SOEs and model enterprises, and 7 SDOs. China Mobile and partners are committed to promoting AI service level, contributing to the digital transformation for thousands of industry.

3.1.2 NTT DOCOMO: Lifestyle Co-Creation Lab enables a well-being society

NTT DOCOMO launched the "Lifestyle Co-Creation Lab" in September 2021 as an initiative to verify the value of technologies that make people's lives richer and more convenient, together with its partners, by leveraging advanced technologies.

Against the backdrop of recent rapid changes in lifestyles, DOCOMO believes that it is important to deeply understand the needs of each individual from a variety of

perspectives across industries, and to create new values that are friendly to people, in order to develop services and products in the future.

Aiming at creating a "Well-being Society" in which all possibilities expand while each individual shines and is close to each other, DOCOMO is working to refine their technologies and create new lifestyles by combining the technologies and assets from DOCOMO and partners across industries and companies.

This initiative will make it easier to verify the value of the AI technologies DOCOMO is developing toward full-scale commercialization. Based on the results, services and products could be smoothly developed to contribute to a better life.

In order to realize the Well-being Society, where people's lives become richer, the following areas of efforts are defined, together with the SDG issues being considered: Safety, Healthy, Life, Working/Learning and Enjoyment. Based on these five concepts, various value verification activities are conducted in DOCOMO .

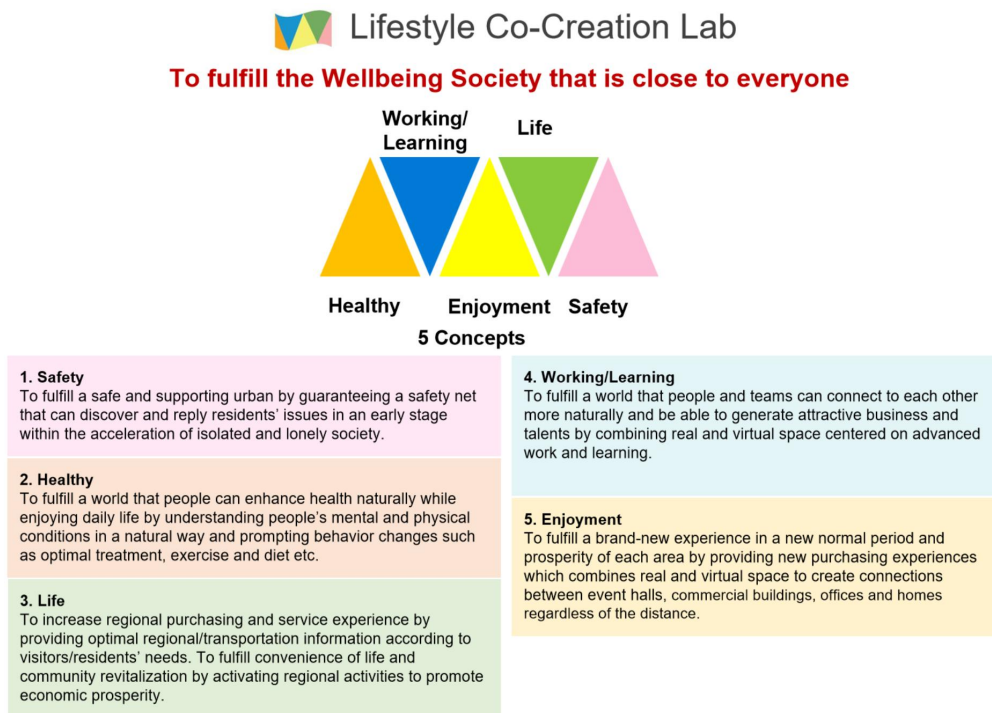


Figure 5 Focus of Lifestyle Co-Creation Lab

In addition, through this initiative, DOCOMO has been developing an Innovation Co-Creation Platform that enables the use of DOCOMO technologies, such as AI and IoT, as well as NTT Laboratories' technologies in various industrial fields. By using this platform, the development and deployment of new services could be accelerated without repeated adjustments for each new technology. The platform is also a part of the NTT Group's "4D Digital Infrastructure", which aims to grasp the precise location

of geographic space and various mobile objects, and realize various future projections based on this understanding. Through this initiative, DOCOMO will continue to work with various partners to create new values with the aim of creating a "Well-being Society" that enriches people's lives.

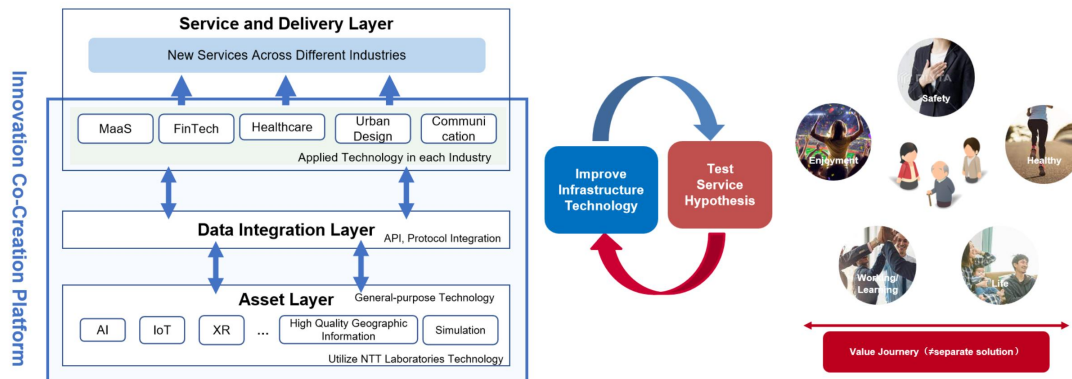


Figure 6 Innovation Co-Creation Platform enables well-being society

1) Lifestyle Co-Creation Lab

The Lifestyle Co-Creation Lab combines multiple technologies and assets from DOCOMO and its partners to verify the value of technologies that make people's lives richer. Through demonstrations at the Lifestyle Co-Creation Lab, an innovation co-creation platform is also provided, enabling the use of DOCOMO technologies, such as AI and IoT, as well as NTT Laboratories' technologies in various industrial fields.

● Roles of the Lifestyle Co-Creation Lab

The following steps are taken, together with DOCOMO's partners, to create new lifestyles.

- ① Drawing up a world view (value hypothesis) that is valuable to people living in the world
- ② Testing/accumulating data by examining the value of services that embody the world view
- ③ Connecting multiple services to maximize user experience value

● Components of Lifestyle Co-Creation Lab

Technology, Partner and Field are elements of the Lifestyle Co-Creation Lab, and value verification is conducted.

Elements of Lifestyle Co-Creation Lab

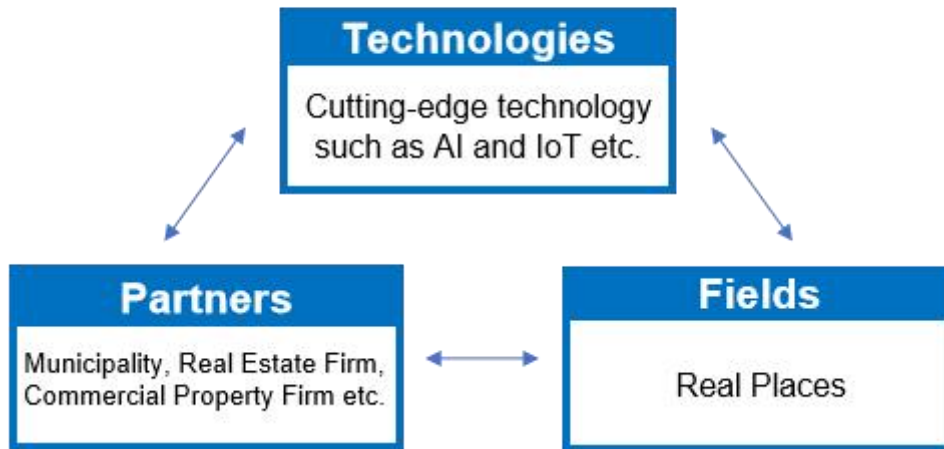


Figure 7 Elements of Lifestyle Co-Creation Lab

2) Innovation Co-Creation Platform

The Innovation Co-Creation Platform makes agile development and delivery of new services a possibility, avoiding the repeated R&D as much as possible, enabling new value creation across industries.

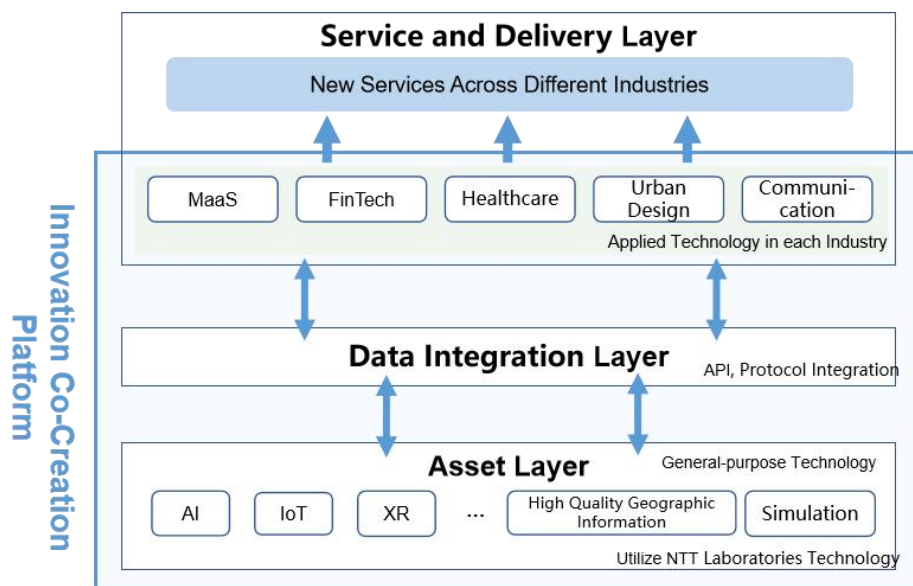


Figure 8 Innovation Co-Creation Platform Architecture

3) Technical Areas to Focus on

- MaaS

Real-time understanding of human flow and traffic information from mobile networks

and car data. Combined with public transportation information, it predicts future congestion and traffic congestion in cyberspace. It helps user stay close and comfortable while optimizing transportation planning. It could also suggest a side trip to user's liking, e.g. if the train is crowded, it will suggest a bus route that is less crowded. It helps user have comfortable travel without congestion, and further solves various problems, empowering sustainable development of the society.

- Healthcare

Predict and visualize user's current and future health in cyberspace from sensors around and smartphone data. By understanding each person's health and preferences, and encouraging them to act healthily with advice like family, this solution helps preventing diseases before they happen. Even when users do get sick, the health data stored on the Internet will be used by medical institutions. Users can easily receive medical care and treatment anytime, anywhere with a deep understanding of themselves.

- Urban Design

For construction sites, data obtained from cameras and sensors, together with topography and climate data, are used to propose safe sites and optimal work plans in cyberspace. DOCOMO is committed to making people who build happy communities excited, secure and happy. In addition, the entire community is reproduced in cyberspace from data obtained during the construction process. Combined with human flow and health data, the solution helps creating a happy and smart city where people continue to connect and shine.

- FinTech

With the expansion of cashless payments, there are many days when people don't take their wallets out even once. Buying goods and services with a smartphone or card means that they are recorded as data at the same time. Being handled properly, they could help people know how they have lived and what they further need to do for the future. FinTech can help users design their life plans and enrich their lives every day.

- Communication

During the COVID-19 pandemic, telecommuting and online meetings spread rapidly. Online communication, on the other hand, has caused problems such as one-way conversations, awkwardness, and a decrease in idle conversations such as small talk. In this initiative, DOCOMO is working to create a casual environment in which speakers and listeners can better communicate with each other through improved UIUX and is changing the way people communicate online.

3.1.3 KT- AI Full Stack, make everyone can easily experience, utilize, share and compare AI.

KT has declared its transformation into a digital platform company “DIGICO” in 2020, and has been accelerating innovation for customers in various business fields such as media, finance, B2B and etc. Based on AI, Bigdata, and Cloud technology(ABC), KT has been leading the digital transformation with digital innovation and ESG. KT has proven this by surpassing \$18 billion in revenue for the first time in 2022, and the company continues to grow steadily.

KT conducts various B2B/B2C businesses with the largest IDC in Korea, cloud service, and 5G public/private network(wired/wireless). In particular, through the partnership, KT has strengthened its R&D capabilities and built an R&D chain “AI Full Stack” from infrastructure to service. By providing compatible AI solutions, KT makes customers and industries be more convenient and efficient, and increase their productivity.

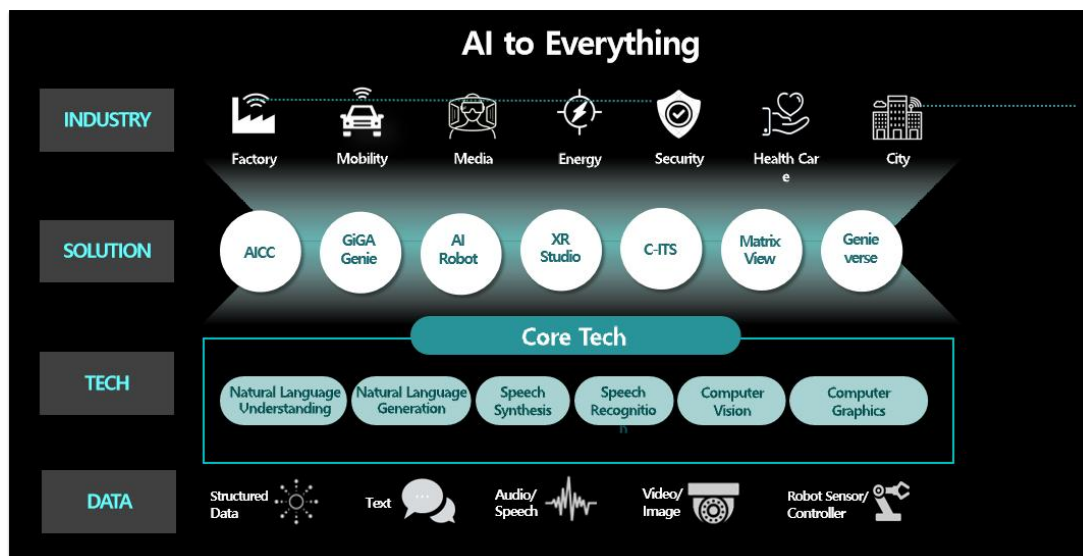


Figure 9 Full Stack Solution applicable to different industries for smart city
 KT promotes business and R&D to develop a goal-oriented AI that is customized to a vertical domain not just general use. Focusing on the target field, we establish an efficient structure and create innovation of services that reflect the needs of the field with “Full Stack” and “Full Packages”.

1) Core Tech for LLM

KT established and implemented “AI Full Stack” strategy to empower AI business so that it built an R&D chain consisting of technologies from HW/SW infrastructure, framework, AI model, to services and applications. We provide our own model which is able to be customized and service that meet the needs of customers with cost-down

reasonable. The compatible E2E services/applications make the users migrate to KT Cloud by providing the full-stack AI solutions on cloud.

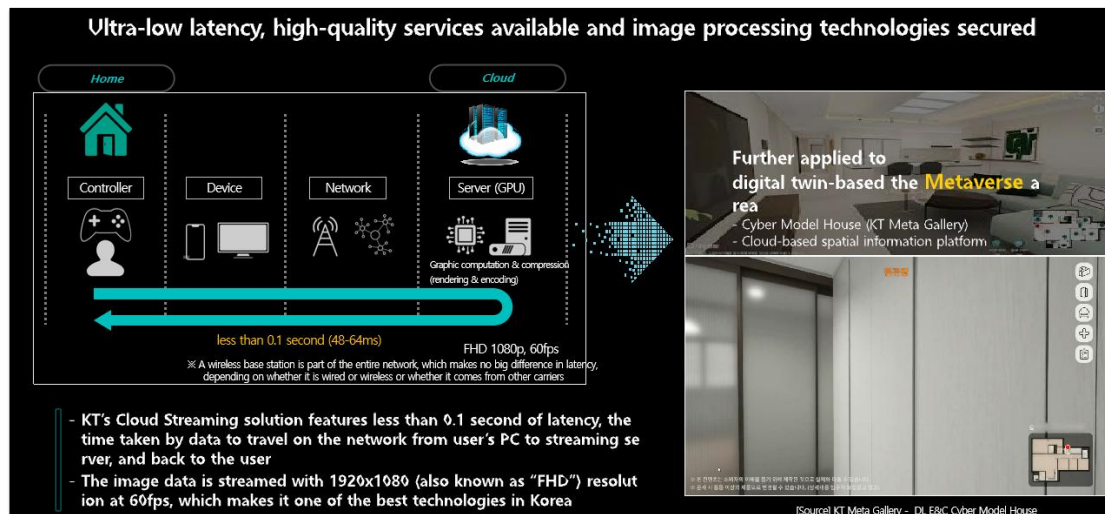


Figure 10 Dedicated Network and Cloud Service

“Large AI Mi:dm” is developed with various ML/DL theories and techniques to achieve the model’s performance in learning and inference. And we also R&D the core techs to provide customized one to each targeted field. We focus on the industry needs such as security, accuracy, retrieval and etc. and reflect them into the model to improve user experience. We strive for a cost-efficient AI model with combined lightweight technology that would be widely applied in the industry.

2) Co-operation for AI semiconductors

When a company uses AI for its business, the high cost would be a problem. KT is working with partners on HW infrastructure R&D for the LLM to mitigate the issue. By cooperating with Rebellions, Fadu, and moreh, we are conducting R&D on AI semiconductors that will reduce the cost of running AI. We also provide SW technology such as HAC(hyperscale AI computing) on cloud service.

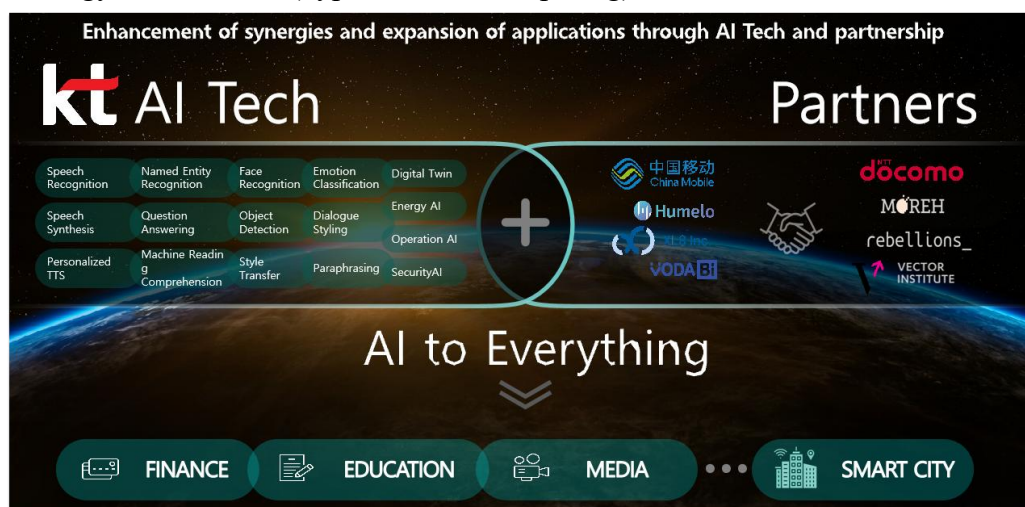


Figure 11 GenieLabs——KT AI Eco Playground

3) The strategy for AI Eco

KT defines the business layers of the AI market and develops an appropriate strategy for each business layer. We offer a range of services based on our technical capabilities and assets such as IaaS(infrastructure), MaaS(foundation model), PaaS(LLM platform), and SaaS(API). We also provide LLM platforms “Mi:dm studio” and “Genielabs” which allow customers and 3rd party to customize the model according to their needs or preferences. Users can easily fine-tune, optimize the model for a domain, and build their own persona on the platforms. They can also get maintenance, updates, and additional features.

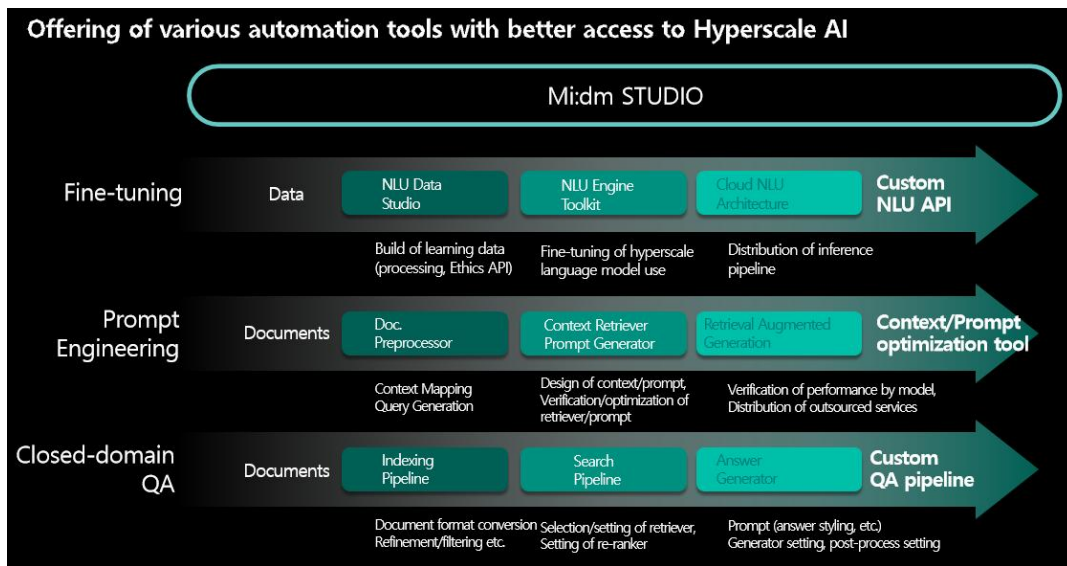


Figure 12 Mi:dm STUDIO Architecture

4) AI Use Case

“Large AI Mi:dm” is used for real industrial services such as AICC, consultation and retrieval, and is expanding its application areas to healthcare, medical, robotics, and traffic management. With additional training, it can also be used in professional fields such as finance and law.

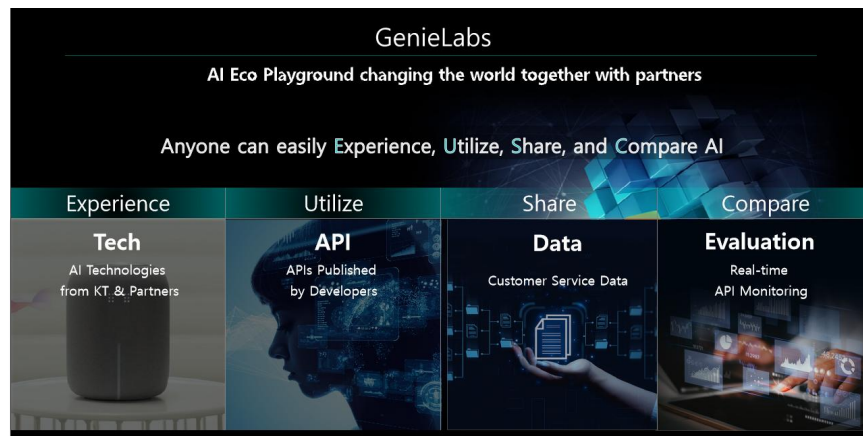


Figure 13 KT GenieLabs

KT provides services based on emotion understanding and empathy technique, such as parenting counselling and senior care.

3.2 Key Enabling Technologies

3.2.1 Intelligent Computing Infrastructure

Intelligent computing infrastructure is mainly composed of computing resources and data resources that support AI model development, training or inference. Computing resources include AI training and inference chips, heterogeneous intelligent computing servers, and intelligent computing centers/ supercomputing centers based on chips and servers. The rapid development of AI brings the rapidly growing demand for computing resources, putting great pressure on storage, computing, bandwidth and security. Therefore, a large number of edge computing nodes will continue to be added between the cloud and the end, thus deriving the computing power and data collaboration of end-edge-cloud. By building a "intelligent network brain" can efficiently optimize the selection and matching of computing nodes, network paths and AI models, enabling the unified scheduling and opening of computing, network and both large and small models, to better support the development, training and inference tasks of AI models. Data resources include basic data collection and labeling, data storage, data governance, data confirmation, etc., as well as data centers built on these. Data centers serve the training and optimization of AI model, cooperating with the database design, data collection, data cleaning, data labeling, data quality inspection and other tasks of data service providers to ensure qualified data governance and security, and finally provides high-quality data sets for AI model training.

3.2.2 AI software stack/architecture

As a software system, artificial intelligence software stack aims to build an end-to-end technology development system of software and hardware integration based on AI chips and other hardware. It includes not only the underlying computing resources and data sets, but also software modules such as algorithm framework and tool chain. These software modules depend on each other in the software stack, carrying data flow, calculation flow and control flow, which can assist in one-stop full-process model development and training, at the same time, improve the adaptability of software and hardware to the maximum efficiency, and increase the efficiency and convenience for developer's model training. At present, the AI software stack is also

accelerating on standardization, unifying the full stack technology system, increasing scalability and integrating various open source models.

3.2.3 large foundation model

Large foundation models (pre-training model) refers to a model with huge parameters trained by self-supervised learning based on a large number of unlabeled data, which can complete a wide range of downstream tasks under the prompt of small samples/zero samples.

Compared with traditional machine learning models, there are following differences: (1) training data: the training data of traditional machine learning model is at hundreds or thousands, which needs manual labeling. The training data required by large models is often over 100 million without manual labeling. (2) parameters: the parameters of the traditional model are small, while the parameters of the large model are usually of billions. (3) Different training periods: the training duration of traditional models is generally at minutes/hours, and the training of large models often takes month long. (4) The universality: traditional models are trained for specific scenes/tasks, while large models are general models for multiple tasks. The success of large foundation models has completely subverted and reshaped the AI technology paradigm. The first generation of AI technology paradigm is small models for single task. Each model can only predict and complete one specific task. The second generation of AI technology paradigm is pre-training and fine-tuning, which needs to manually label a certain amount of data and fine tuning of neural network parameters for specific downstream tasks. The third generation takes a large foundation model with a huge number of parameters as the base. It can complete various tasks such as abstract, question and answer, dialogue, writing, etc. under the prompt of small samples or even zero samples.

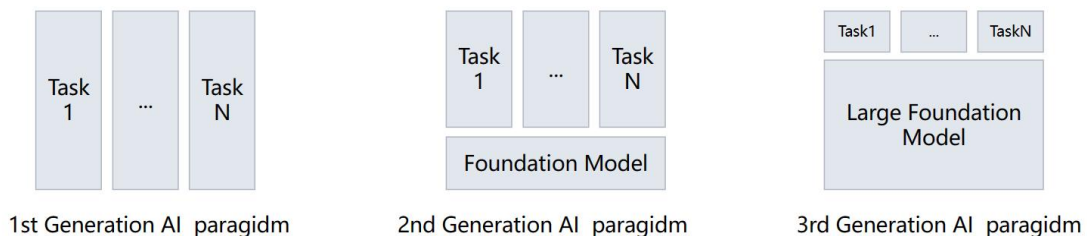


Figure 14 Evolution of AI paradigm

3.2.4 AI operation

With the rapid development of AI industry, more and more machine learning is applied to various services of terminal cloud. The end-to-end process of machine learning involves data exploration and analysis, model development, model training, model evaluation, model deployment and effect evaluation. MLOps was first proposed by Google in 2015 to solve the technology stack problem of machine learning, aiming at unifying machine learning system development (Dev) and machine learning system operation (Ops), enabling an efficient ML system construction process through automation and monitoring of integration, testing, release, deployment and infrastructure management. A MLOps-based machine learning project life-cycle is as follows:

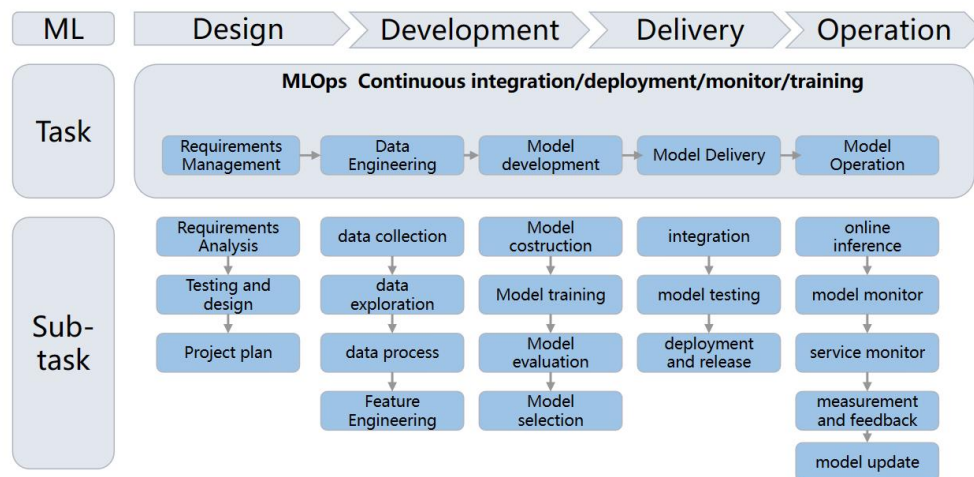


Figure 15 MLOps Framework

So far, MLOps has been widely used in IT, finance, e-commerce, manufacturing, etc., helping service operation, product development, marketing, risk prediction and supply chain management. For CSPs, due to the huge number of users, in order to ensure the stability of the model, the operation monitoring after the model launch is particularly important. Taking the application scenario of network intelligent transformation as an example, MLOps is considered key technology for evolving into L4 highly autonomous network³. It can be used to manage the R&D assets of intelligent network applications such as data, models and algorithms, realize large-scale automatic deployment of intelligent network applications, continuously monitor the deployment effect, avoid the risk of model degradation, support the continuous optimization and iterative training of models, and standardize the workflow of applications from model development to delivery and operation.

³ TM forum. 《Autonomous network white paper v4.0》 2022.09

With the rise of large models, LMOps and LLMOps are also being discussed, that is, "MLOps for large models/large language models", which inherits the overall framework of MLOps and the whole life cycle of machine learning, and makes micro-adaptation for the changes of large models in each link, thus solving the challenges of introducing large model-driven applications in production environments. For example, high-quality small and medium-sized data sets should be produced through data engineering as input for model development and optimization according to the needs of specific industries, fields or scenes, including data collection, labeling, processing and generation, etc., but excluding feature engineering. In the link of model optimization, the parameters of large model are adjusted or redeveloped through various optimization techniques to achieve in-depth and refined model optimization, so as to adapt to specific needs, including prompt engineering, model retraining, model fine-tuning and RLHF. In the model delivery stage, the optimized model forms a deployable model or service, including model compression/transformation, testing, service deployment, etc..

3.2.5 AI security

AI security refers to taking measures to protect the hardware, software and data of AI system and the services based on those, so that they will not be accessed, leaked, destroyed, modified, reviewed, checked, recorded or destroyed due to accidental or malicious reasons, thus ensuring the confidentiality, integrity and availability of information. On the one hand, generative AI-related technologies have problems such as data leakage, false information and pornographic and violent content. On the other hand, with the emergence of various generative AI applications, the amount of data, software and hardware applications, etc. will increase substantially, and the attack scope will increase accordingly. Generative AI will also greatly reduce the threshold of hacker attacks, and generate code such as vulnerability attacks, phishing attacks and harpoon attacks in batches, and the amount of attacks will increase. AI technology development will bring unprecedented challenges to network security, showing a new trend of "large-scale intelligent computing is widely used in the security field, and attack and defense has entered the era of intelligent confrontation".⁴

Generative AI will change information security industry into a different development logic, so will be the security products and services. The "AI+ security" model will broaden the development space of the whole industry in the future. A security

⁴ Internet+Development Association of China. 《10 trends in network security 2023》 2023.03

management platform with "AI security brain" as the core is expected, promoting the integration of hardware products, software products and security services.

3.3 Summary

Digital economy is a common opportunity for all industries, especially for CSPs. 5G and AI are considered as two "cornerstones" for the development of traditional industries in the future. Through better network services, more efficient operation and business efficiency, and technology innovation with AI as the core, operators have gradually become the major force of developing digital economy and leading industry transformation.

The investment and development progress of telecom AI in China, Japan and Korea has been at the forefront and constantly strengthening. CSPs share some very same interests including, attaching great importance to the AI platform and AI infrastructure development, gathering excellent AI technologies through both self-development and industrial collaboration, strengthening the industry standards and ecology building, accelerating the innovation and commercialization of ToB AI applications, and providing secure and controllable services, thus helping social and economic development and improving people's lives.

4. Best Practices of CSP

4.1 China Mobile Practices and Use Cases

China Mobile takes AI as a strategic technology innovation. By integrating computing, network and intelligence into a new type of infrastructure, China Mobile has built a comprehensive intelligent service capability.

4.1.1 Network intelligent transformation, towards L4 highly Autonomous network

China Mobile is the first CSP in the world to propose the goal of “reaching L4 network-wide autonomy by 2025”. To better introduce AI into world’s largest and most complex network, China Mobile has taken an approach of platform development, and ability and application innovation.

- Intelligent Network simulation platform

The intelligent network simulation platform provides a 3D twin verification environment for network AI ability incubation and model training, especially for decision intelligence abilities that need frequent adjustment and observation of the network for continuous learning. The platform selects multiple types of existing network areas to build end-to-end full-range elements intelligent simulation abilities driven by real network data, including: intelligent user behavior and trajectory simulation based on urban maps, real environment coverage and full-stack simulation of base stations, flexible core network and business service simulation, flexible network fault injection and cross-layer data collection, etc., which promotes technology innovation breakthroughs in perception intelligence, analysis and decision intelligence, to better incubate and verify newly developed AI abilities.



Figure 16 3D intelligent network twin

- Intelligent Network ability platform

It provides network-wide ability training, deployment, ability management and operation services for the large-scale application of network AI capabilities, and accelerates the closed-loop efficiency from R&D to deployment to iterative operation. So far, more than 60 internal and external R&D units have joined the platform, with more than 3,600 network AI abilities, including over 110 high-value abilities that have been reused in more than 10 provinces. The platform currently supports 32 units to make API calls of over 3,000 applications, with a cumulative API call over 2.3 trillion times.

- Intelligent Network Application Innovation

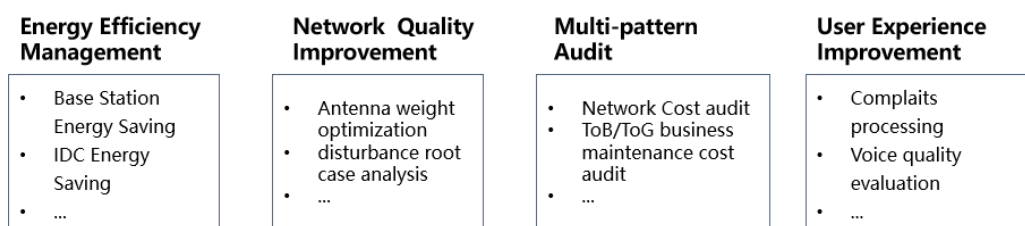


Figure 17 CMCC Network AI Applications

Applications are created around the whole life cycle of network of "network planning, deployment, maintenance, optimization and operation", including 4 major types: energy efficiency management, network quality improvement, multi-pattern audit and user experience improvement, which have been implemented in 31 provinces and achieved remarkable results.

4.1.2 Customer service large model enables transformation towards new user experiences

In July 2023, China Mobile released JIUTIAN customer service LM, which is based on the massive customer service data, business knowledge and experience. It can not only provide services to users independently, but cooperate with human customer service attendants to improve the work efficiency. It will redefine the connotation and mode of customer service to a certain extent, and bring users a brand-new experience.

3 major features of JIUTIAN customer service LM are:

- Connection of the broad scenario and whole process. Based on China Mobile's long-term accumulated business knowledge, interaction and feedback logs, system interface data and expert experience in the customer service domain, the workflow is redesigned for better collaboration between large models and human customer service attendants, and connection of the different scenarios and processes.

- Customer caring-driven. The model is trained based on China Mobile's data set of real people dialogue, which is also the biggest dialogue data set in the industry. The goal is to enhance emotion, empathy and improve user satisfaction, achieving customer caring.
- The three-way communication field of "large model-human customer service attendant-user" with man-machine cooperation. For example, after knowing that the user can't watch TV, the large model can detect the failure of user's home broadband service and guide the user to find the root cause automatically. And if the root cause were too complicated, it can actively connect the human customer service attendant. The human attendant can quickly take over the conversation, continue to investigate the problem, and wake up the large model again when necessary. They are seamlessly connected and complement each other.

4.1.3 Digital government: Government affair LM provides high-quality public services and empowers modernized social governance.

China Mobile has launched JIUTIAN HAISUAN government affair LM, along with a number of AI abilities, to help building the AI engine of digital government, empowering key applications, providing high-quality public services to citizens, and effectively improving the digital intelligence level of social governance.

- **Use Case 1: The first AI platform for digital government in China.**

The AI platform of Heilongjiang digital government project is designed and developed around the concept of "3+1+1+N": 3 foundation platforms of inference, training, and video cloud, 1 atomized AI ability center, 1 application support platform, supporting N upper-level applications.

The platform connects multi-source heterogeneous data such as video, image, voice and text, and provides intelligent analysis services for applications including smart government affairs, precise governance and public services. By innovating means of supervision, improving workflow efficiency and creating a convenient environment, it fully solves the challenges at different levels of provinces, cities, departments and bureaus, accelerates the digital government development.

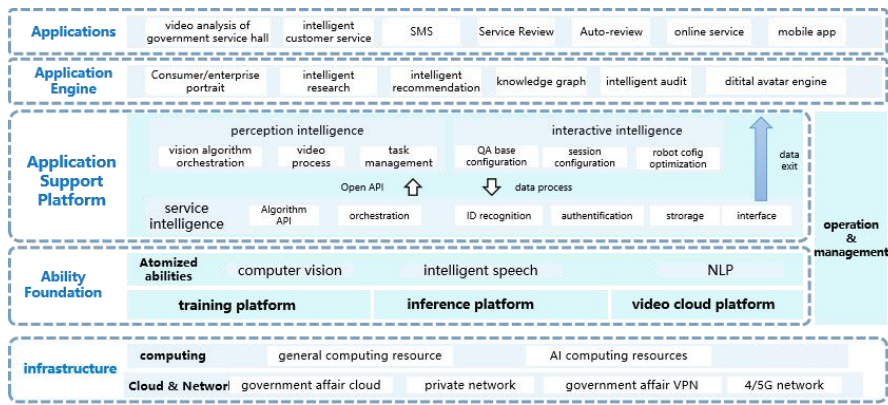


Figure 18 ToG AI platform architecture

● **Use Case 2: Smart customer service for government affairs**

Project team has leveraged inference platform, NLP abilities and intelligent speech abilities to build a government knowledge graph with 200,000 entities, 10 million business-related associations, and 1 million standard question and answers to continuously optimize the response speed and accuracy of customer service. China Mobile has also created the leading intelligent guidance function, providing a more convenient, intimate and intelligent one-stop customer service experience for both enterprise and individual users. The project is fully implemented in Gansu province, empowering a number of provincial and municipal government systems, serving 25 million citizens.

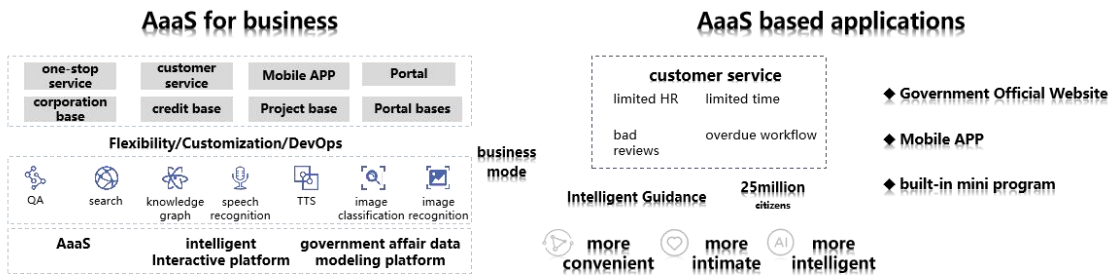


Figure 19 Smart Customer Service for Government Affairs

4.1.4 Promoting AI Talent Development : AI Learning and Practice Platform

Based on China Mobile's cloud network resources and JIUTIAN product system, Bisheng platform is developed to provide AI learning and practice service for universities and colleges, including course development, computing center construction, entrepreneurship cultivation, talent training and other services. By enriching AI teaching and research data resources and innovating AI talent training mode, Bisheng currently serves 160,000 teachers and students in over 200 universities and

colleges.

Taking Hangzhou Dianzi University as an example, a "5G+ Intelligent Teaching" solution is created with a one-stop AI teaching management platform to meet the needs of online teaching and cross-regional interaction. A JIUTIAN AI Collaborative Innovation Center is also established to carry out in-depth cooperation in scientific research, teaching, internship and practices, etc. So far, a number of AI training courses, as well as collaboration projects have been jointly developed and launched on the platform. Multiple AI innovation and entrepreneurship practice bases are established to provide students with exclusive practice scenarios, leave them with unlimited imagination to turn textbook knowledge into practical applications. A series of activities are organized such as AI contests among different universities, to provide students with a global vision and competitiveness.

4.1.5 Smart city: improving city management efficiency

China Mobile has developed JIUTIAN city AI platform, carrying multiple intelligent solutions for “bright kitchen”, safety prevention and control, and scenic spot management with the intelligent analysis ability of video stream as the core, promoting intelligent city management.

● Use case 1: Smart city creation in Yunnan Province

The smart city project in Chuxiong Prefecture is consisted of 4 parts: infrastructure, digital platform, N applications and 1 intelligent operation center. The infrastructure mainly includes the data acquisition infrastructure of video, temperature and humidity and other IOT sensing devices. Digital platform refers to city data platform, open API platform, cloud service, etc. Relying on 5G, N applications are created such as 5G+ smart sanitation, 5G+ smart urban management, 5G+ smart parking, 5G+ farmer's market, etc. An IOC supports 4 major functions: operation monitoring and analysis, event management, emergency correlation and decision-making, supporting city governance in an all-round way and providing one-stop decision-making command service for the leadership. A digital twin city is constructed by integrating the data from various commissions and bureaus, realizing the comprehensive display, analysis and prediction of people and things, enabling the fine management.

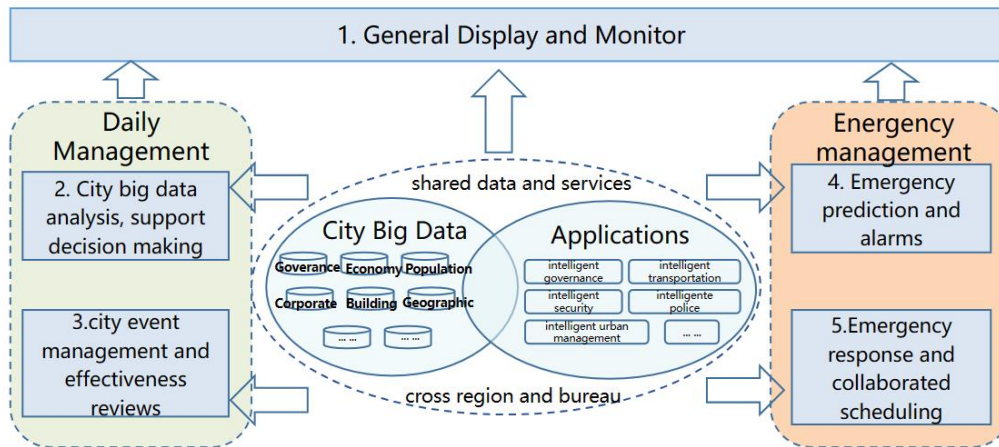


Figure 20 Smart City Solution

4.1.6 Smart Manufacturing: Helping Safe and Efficient Production and Realizing New Industrialization

China Mobile provides AI-assisted manufacturing and management services for industry customers in different domains, helping industry customers to reduce costs, improve efficiency and improve production environment.

- **Use Case 1: intelligent detection of cracks and depressions in aircraft skin for China Eastern Airlines.**

Airline companies are facing the pain points of lacking professional maintenance personnel, high maintenance cost and safety guarantee. Combing 5G, cloud computing and digital twin technology, the cross-interface, cross-protocol and cross-platform connection is realized for seamless integrating and interaction between physical and virtual twin world . Different aviation maintenance scenarios are reproduced on the digital twin platform, optimizing the maintenance workflow, and developing solutions such as 5G+AR assisted maintenance, 5G+VR maintenance training, 5G+AI assisted detection and 5G+ connected airbase in combination with hardware such as inspection AGV, AR glasses, UAV, etc. Intelligent inspection replaces traditional manual inspection to improve efficiency and quality. The defects such as dents on the aircraft surface can be effectively identified, and the preventive maintenance of key aircraft parts can be realized, which can further reduce the maintenance cost, guarantee the safety.



Figure 21 Surface detection example

- **Use case 2: industry quality inspection, improving production efficiency for edge manufacturing**

Quality inspection is a key link in manufacturing and the last defense line of product quality. With the continuous upgrading and evolving of manufacturing, traditional quality inspection methods become unable to meet the requirements of high efficiency and high precision. Relying on IoT, AI, edge computing, China Mobile has created intelligent industry quality inspection products with 3 major advantages:

- a. combination of self-developed detection algorithms and optical automation environment.
- b. Industry-specific middle-sized model to reduces retraining time after product update and replacement.
- c. The inference optimization based on a variety of heterogeneous computing force, with efficient, intensive and low-cost deployment.

So far, China Mobile's industry quality inspection products have been promoted in the automotive industry, forming a strategic cooperative relationship with CRRC, FAW, SAIC and other China's moto industry partners. The products are aimed at the quality inspection of core parts such as automobile engines, wheels and turbochargers, achieving a waste detection rate of 100% and a false detection rate of less than 1%, improving the detection accuracy and reducing labor costs.

4.2 NTT DOCOMO Practices and Use Cases

Through "Lifestyle Co-Creation Lab", DOCOMO is promoting to verify the value of technologies that make people's lives richer by leveraging technologies and assets that owned by DOCOMO and its partners. Some specific examples of demonstration experiments and commercialization will be introduced as follows.

4.2.1 Patrolling System Using Image Recognition AI

NTT DOCOMO(DOCOMO) and NTT Communications Corporation (NTT Com) have succeeded in a demonstration experiment, building a system that enables users to virtually patrol construction sites from remote by using image recognition AI. The system detects materials that inhibit on-site work, and shares instructions for moving those materials with workers. With this system, DOCOMO and NTT Com aim to reduce the burden of on-site patrol and long working hours for workers, and the goal is to commercialize this system in 2024.

By inputting information such as materials, working spaces and entrances in advance, the system is able to detect materials that inhibit work in areas of the construction site with the use of image recognition AI, and also provides workers with instructions for handling those materials. Specifically, the system analyzes pictures taken by a fixed camera every five minutes with image recognition AI to determine whether any material inhibits work, and gives instructions such as locations to remove those obstructs to workers if necessary. This enables more smooth and efficient communication between workers at construction sites.

In the demonstration experiment, cameras were installed in front of the entrance and the elevator at the construction site, and obstructs were successfully detected using the system far from the site. Moreover, instructions were made for each detected obstruct, and were given to workers to take actions. The result proves the effectiveness of this technology: among the obstructs detected by the system, the probability that they were actual materials(precision) was over 90%, and among the actual materials, the probability that the system detected them as materials(recall) was over 80%⁵. It is the first time in Japan⁶ to establish a technology to detect materials that interfere with works at construction site from remote.



Figure 22 Patrolling System View

⁵ Accuracy of evaluating all materials.

⁶ As of 2023/08/07, according to DOCOMO.

DOCOMO and NTT Com will upgrade this system and aim to put it into practical use in 2024 as a new function of “Dan-8 Area”⁷ from “DOCOMO Construction Site IoT Solution”⁸ provided by NTT Com. Through this initiative, DOCOMO and NTT Com attempt to reduce the burden of on-site patrol, to contribute to the efficiency of construction work and to solve the problems such as workers’ safety at the construction sites. In addition, DOCOMO will utilize digital twin computing⁹ to realize a digital twin on site and carve out a smarter future.

4.2.2 AI for estimating frailty in the elderly

Frailty is a stage before the need for long-term care in which mental and physical functions become weaker with age. In recent years, there has been an increasing need for early detection and intervention of frailty in order to lower nursing costs. However, it is difficult to conduct wide research and respond to the needs by using existing methods such as face-to-face examinations and questionnaires.

AI for estimating frailty in the elderly is a solution with the aim of health management. It can easily visualize frailty risk and suggest lifestyle changes by utilizing smartphones, which have become popular among the elderly.

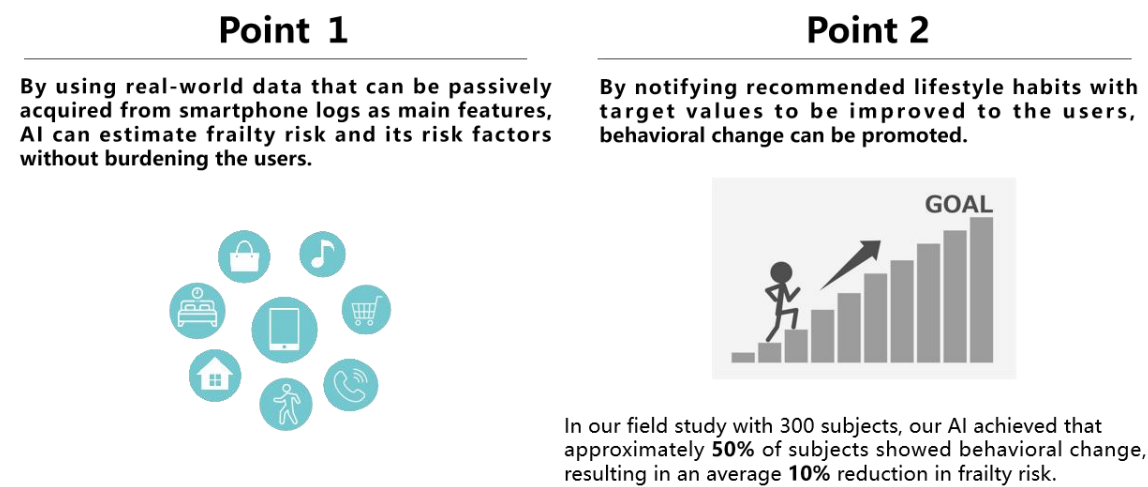


Figure 23 Major features of frailty estimation AI

⁷ Dan-8 Area is a solution with the aim to improve the efficiency of setups and operations at construction sites. It enables users to complete operations such as equipment adjustments through an application..

https://www.ntt.com/business/services/iot_construction/dan8_area.html

⁸ DOCOMO Construction Site IoT Solution is a solution that supports safety, efficiency and information sharing at construction sites.

⁹ Digital twin computing means building a twin world of the real world within digital spaces by expressing objects or human beings in a digital way.

In addition to providing frailty estimation AI, DOCOMO has also developed a HealthTech platform that integrates and incorporates various types of AI related to healthcare. By linking the platform with various types of API, solutions including frailty estimation AI can be used in a wide range of industries. In the future, DOCOMO will develop more AI solutions to realize a world where people can naturally become healthy in their daily lives.

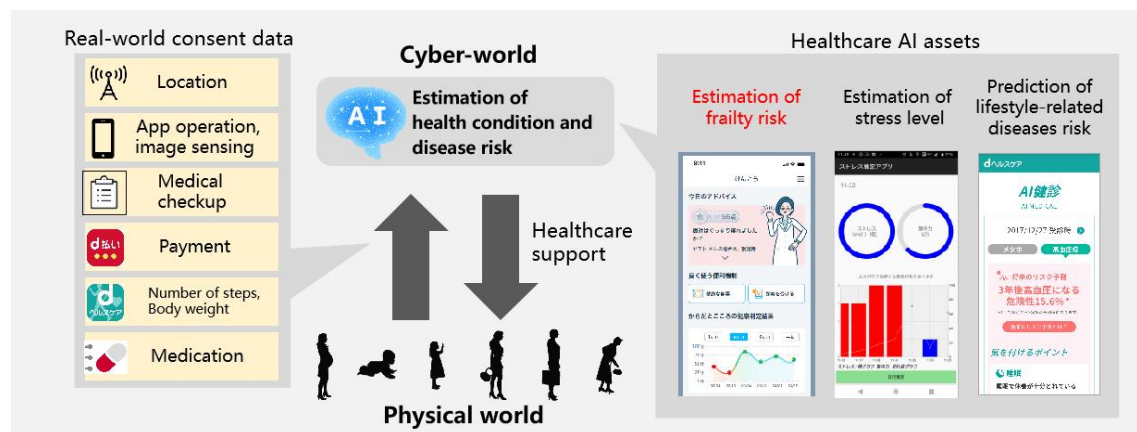


Figure 24 DOCOMO HealthTech Platform

4.2.3 River Watch AI

In recent years, torrential rains and linear precipitation zones have caused frequent floods all over the country. As a result, it is required to detect and respond to floods, and to reduce the damage caused by it in advance in order to realize safe and secure smart cities.

Combined with AI technology from NTT Comware, DOCOMO has developed a river monitoring system with relocatable mobile cameras, which is cheaper and easier to setup than former monitoring cameras. And this system does not require manpower to monitor either.

The mobile cameras are battery-powered so they can be easily set at any places. Moreover, because rise of water level can be automatically detected from the camera images with AI technology, there is no need for constant human surveillance, which can significantly reduce monitoring cost. In combination with the forecast information provided by Japan Meteorological Agency, the water level can be predicted a few hours ahead and actions can be taken to deal with floods.

This system can be used for situations such as monitoring during the rainy season, monitoring at places that have been damaged by floods, river research before construction etc.

Mobile cameras were installed in several rivers in Kobe City to carry out the

demonstration experiments. Data such as river images, rainfall information and water level has been gathered during the experiments. The AI technology that is able to determine the current water level based on the camera images, and to predict future changes of water level based on rainfall information and current water level has been developed. Results of the experiments show that only by installing a mobile camera, the peak of the water rise can be accurately predicted within an error of 10 minutes. Also, the water level can be predicted within an error of 16% of the average. The trial solution has been provided to corporations and municipalities across Japan since 2023/6/28 with the aim of improving AI functions and usability of the system. And the goal is to launch this service before 2024/03.

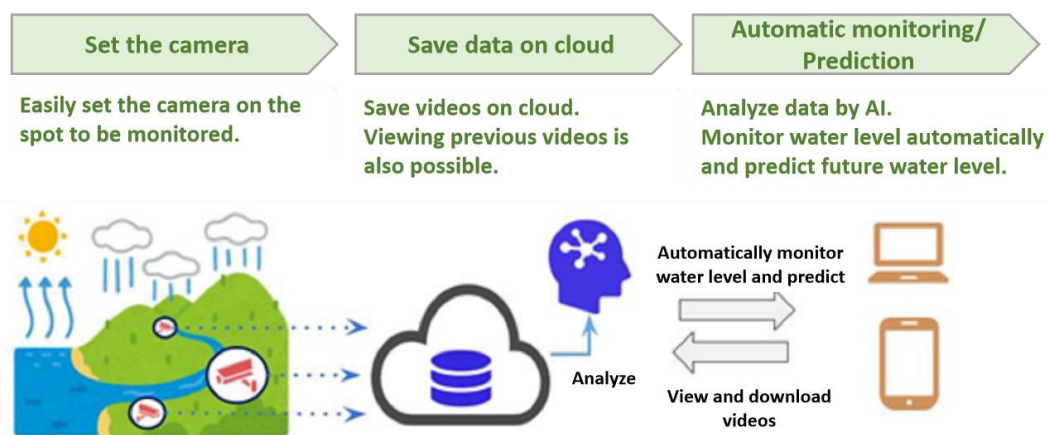


Figure 25 The River Watch AI Framework

4.2.4 Customer Behavior Estimation AI

In Japan, the revised Personal Information Protection Law came into effect in April 2022, and businesses are increasingly being asked to balance information protection and data utilization. To securely manage personal information, NTT DOCOMO utilizes DOCOMO MEC which provides a closed network. By automatically and accurately erasing personal information from camera images, customers' behaviors and interests can be visualized using AI deployed on the MEC. Demonstration experiments have conducted to create new values in stores and offices.

With this technology, necessary data such as a customer's foot traffic, tactile interactions, and gaze can be extracted from the video, providing insight into their behaviors and interests, after personal information such as face is automatically deleted for privacy protection. Also, the background of the deleted person can be automatically changed to make the image more natural than the ones processed by mosaics. The images before the application of this technology are stored in a closed network for security. It is transmitted to the DOCOMO MEC network and processed

in the DOCOMO MEC network to further reduce the risk of information leakage in case of a cyber attack.

In addition, images taken with a smartphone or surveillance camera can be used to detect the presence of a person at a third level that is automatically created by acquiring and projecting information of the time of presence, eye gaze, and object touch to a virtual space by using a pre-existing 3D model of the space. For example, this technology can be used at convenience stores and retail establishments to increase the number of customers. In addition to counting, gender and age estimation, customers' touch and eye of products, in-store traffic lines, and stop heat maps can be further obtained. With user behavior analysis, product placement changes and improved marketing can be made to improve human flow and enable business growth. As the concept of metaverse keeps attracting attention in recent years, the behavior and interest analysis AI can be deployed in multiple domains and enable different new services together with other VR, AR technologies. Heat maps of walking and stopping points are acquired for analysis of human flow and behavior. The effectiveness of the solution is well verified.

In recent years, business decisions based on customer big data analysis have become more and more important for enterprises, and this technology reduces the obstacles faced by the application of video analysis solutions in terms of personal information protection, and helps enterprises choose the right solutions and make better use of data.

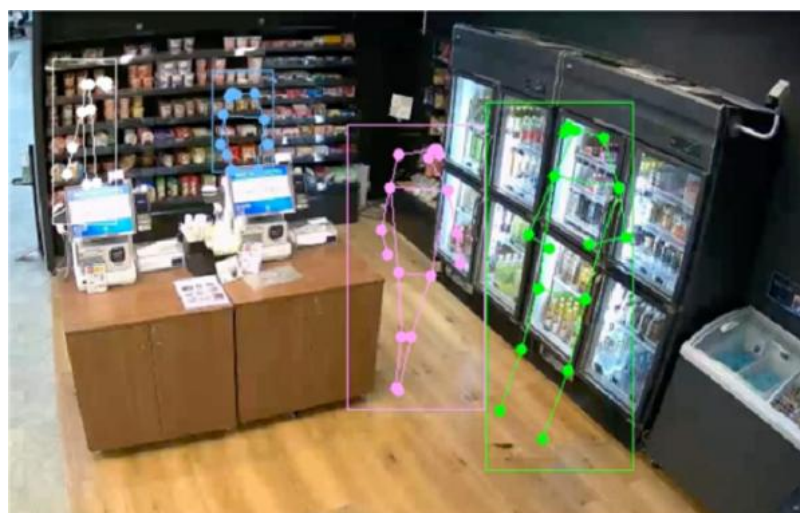


Figure 26 Example of a redrawn customer's behavior without personal information

4.2.5 AI for mobility sharing service

Micromobility-sharing services are services in which users can freely rent and return micromobility vehicles at stations located throughout the city. While an increase in users is expected, the challenge lies in the increasing burden of the "reallocation task," which involves proper reallocation of vehicles at each station and proper maintenance of battery levels due to an increase in the number of stations and micromobility available.

In this system, AI generates an optimized reallocation route and recommends it to the maintenance personnel responsible for the reallocation task. Based on the demand forecast from machine learning, the simulator replicates the movement of vehicles and calculates the estimated number of available rental vehicles and battery-depleted vehicles at each station. The system then generates an optimized reallocation plan by considering factors such as the order in which each station is serviced, the number of collected and placed vehicles, and the number of battery swaps required. Maintenance personnel can then use tablets or other mobile devices to view precisely which vehicles need to be trucked to other stations and which batteries need to be replaced for maximum operation efficiency. This system's unique feature is its ability to generate a flexible and adaptable reallocation plan based on factors such as rental and return history, weather data, date and time, as well as current operational information such as the number of vehicles and batteries per truck, the maximum number of loadable vehicles per truck, and the distance between stations. With the introduction of this system, even inexperienced maintenance personnel can carry out these tasks with the same knowledge and expertise as experienced ones. Moreover, this system is expected to contribute to efficiency in new territories.

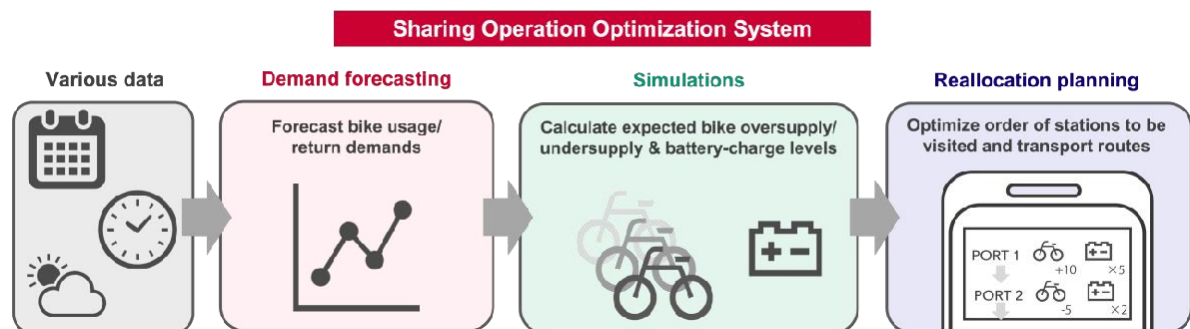


Figure 27 Overview of the system

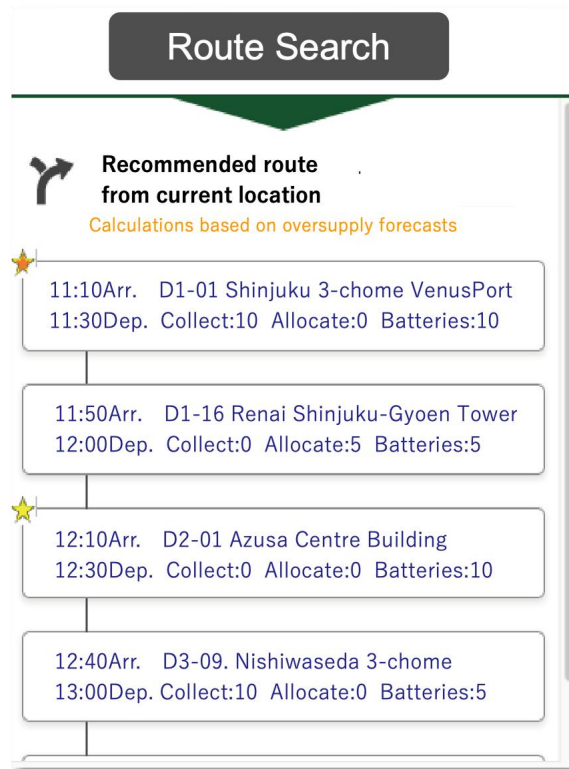


Figure 28 Maintenance personnel's view on reallocation plan(example)

The sharing operation optimization system consists of three technologies: demand forecasting, simulation, and reallocation planning.

Demand forecasting technology predicts the number of rental and return micromobility vehicles per hour for up to 24 hours at each station. By using real-time and statistical data on sharing service usage, along with various data such as weather forecasts, date and time, accurate demand forecasting can be achieved, which fluctuates daily.

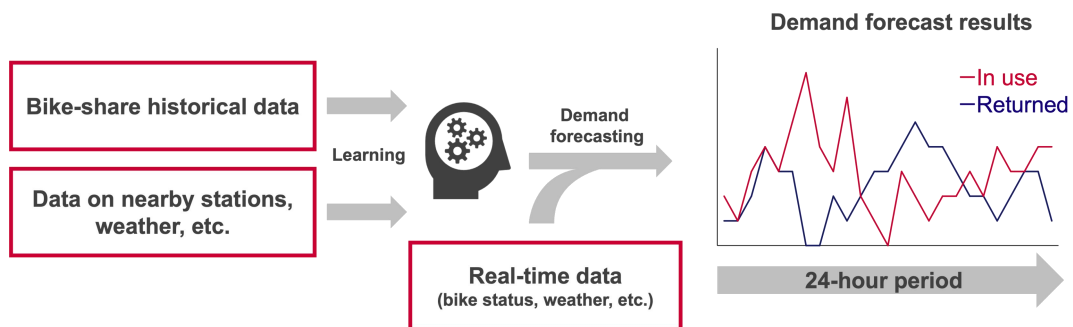


Figure 29 Demand forecasting

Simulation technology reproduces the movement of vehicles from station to station in detail on a multi-agent simulation that expresses the behavior of rental and reallocation trucks. With input values such as real-time information on each vehicle and statistics on movement probability between stations, prediction of the number of

batteries and vehicles at each station is possible. The simulation is updated every 10 minutes, so the latest simulation results can be used to generate the relocation plan.

Reallocation planning technology generates a reallocation route plan that optimizes both the collection and placement of vehicles and the reallocation of batteries, based on the simulation results. This allows for the optimization of the number of vehicles and batteries at each station, as well as prioritizing battery reallocation at stations with high usage rates. In addition, it enables the generation of a reallocation plan to address future problems, such as collecting vehicles from stations with increased return demands. Therefore, this technology is expected to support inexperienced workers and contribute to the efficient development of work routes in new areas.

This system was adopted on April 2024 by DOCOMO BIKE SHARE, INC., the provider of a bicycle-sharing service. Hereafter, the system will be gradually deployed not only in Japan but also to the world.

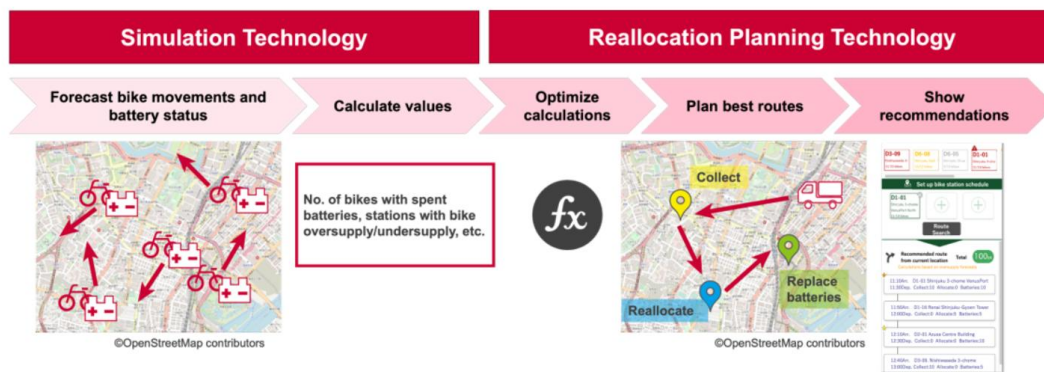


Figure 30 Reallocation Planning

4.3 KT Practices and Use Cases

4.3.1 Media

Since launching the world's first AI TV, GiGA Genie, in 2017, KT has successfully positioned itself as a B2C digital platform company by securing the largest customer base in the domestic smart speaker market with 3.5 million subscribers. This achievement has led to continuous improvements in AI performance. To lead transformations in customers' lives and innovations across various industries, KT has collaborated with numerous domestic and international partners to build an AI ecosystem. Through these collaborations, KT has provided over 120 services, enhancing the convenience of customers' living spaces. Notably, by packaging core

technologies of GiGA Genie, including voice recognition, voice synthesis, natural language processing, and vision AI, KT has contributed to innovations in various B2C and B2B industry sectors and bridging digital information gaps with services such as AI Call Assistant, AICC, AI Space, AI Robot, AI Codiny, and AI Care.

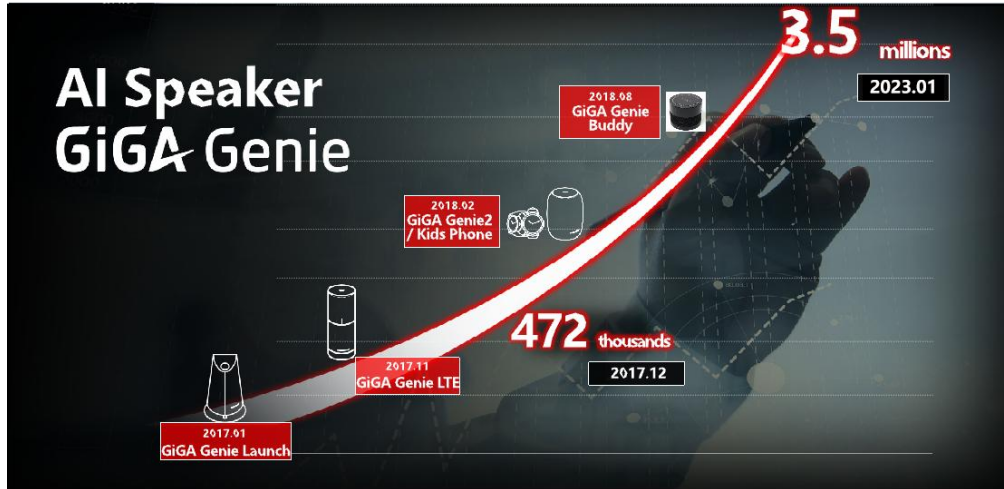


Figure 31 KT GiGA Genie User Growth

4.3.2 Finance

KT, leveraging AI technology proven in GiGA Genie, the largest domestic AI speaker, is innovating the entire Value Chain of the financial industry through the fusion of financial data. This innovation is backed by "Mi:dm," a super-scale AI model specialized in Korean that aims to understand emotions and empathize with humans. The main technologies include:

Knowledge management technology based on natural language comprehension search, which provides detailed content and links to source documents for the vast internal documents and operation manuals.

Advanced digital work environment technology combined with business collaboration tools for tasks such as meeting minutes transcription, summarization and highlighting, and insight extraction.

Enhanced customer service assist technology that offers product query responses and comparison/recommendation customer interactions, bolstered by features like consultation summarization.

Lastly, technology for automatic generation of personalized marketing messages, and drafting market analysis reports and product descriptions.



Figure 32 KT Large AI Mi:dm

4.3.3 AI Contact Center

KT's customer center is leading a paradigm shift in customer service through the integration of AI technology and a customer-centric approach. With the AI Consultation Assist solution, which aids in accurate task handling by freeing consultants from repetitive tasks, features such as automatic classification of VOC types, guidance on customer inquiry-related knowledge, and real-time automatic summary of consultation content have been introduced. As a result, the time to complete customer requests has been reduced from 20 seconds to 5 seconds.

Additionally, in April 2021, a voice virtual consultation was launched, allowing the AI voice bot to directly address and resolve customer issues 24/7 across 70 fields and 650,000 inquiries. This has enabled consultants to move away from mundane repetitive tasks and focus more on empathizing with customers. KT is committed to leading the pursuit of changes and happiness in customers' lives through thoughtful technology and services.

Furthermore, in October 2021, KT launched the mini customer service AI Call Assistant for small business owners who often miss phone inquiries or requests due to tasks like in-store customer service. This AI not only assists with greetings but also provides information on restaurant locations, business hours, parking guidance, and even manages reservations and orders. Since its launch, it has handled over 30 million customer inquiry calls. As of March 2023, approximately 32,000 subscribers are using the AI Call Assistant, ensuring that no valuable customer calls are missed by managing all incoming calls 24/7 and contributing to the increased sales of small

businesses.

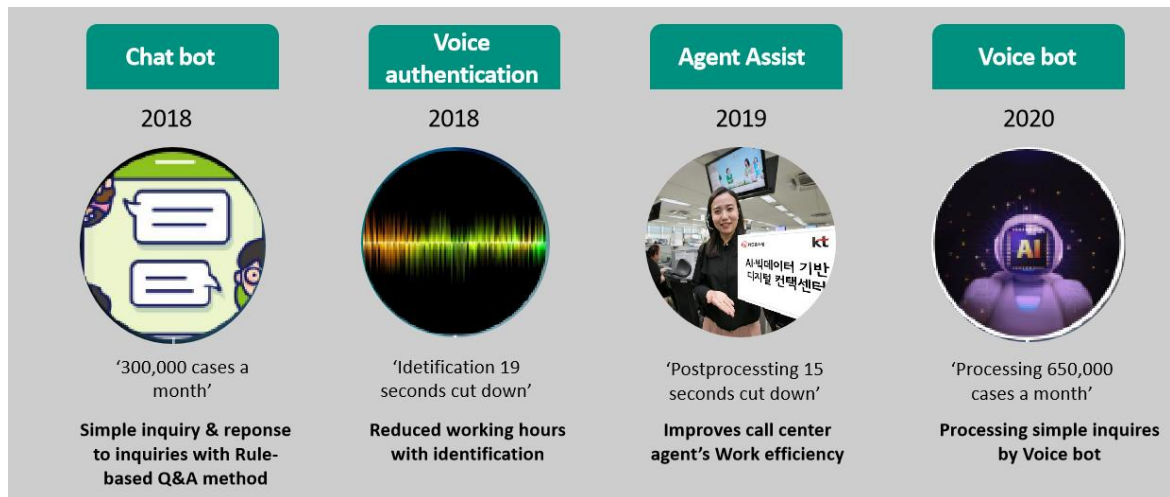


Figure 33 Overview of AICC use cases

4.3.4 Robot

The KT AI Robot is a service robot that combines autonomous driving, robotics, and communication technologies to offer automated and unmanned services. Notably, there's the AI Serving Robot that assists with the strenuous and tedious serving tasks in F&B establishments; the AI Delivery Robot that transports various items in apartments, campgrounds, hotels, hospitals, and offices; and the AI Sanitation Robot leading a new paradigm in daily safety by disinfecting and eliminating viruses.

KT proposes robots tailored to the customer's intended use and provides space consulting, including network considerations, to establish the optimal robot operating environment. Additionally, through AI control, any anomalies in the robot's function are detected in advance and promptly addressed.



Figure 34 KT Robot Example

4.3.5 Education

In alignment with the South Korean educational authorities' efforts to transition towards future education, KT will implement a customized AI Future Education Platform in public education settings this September.

KT's "AI Future Education Platform" is a service that offers tailored teaching/learning solutions. It diagnoses students' learning levels through an analysis of their AI learning achievements and assists them in autonomous learning with customized recommended learning content. Moreover, features such as automated lesson planning, AI-customized assignment creation, and automatic analysis of class/student learning levels reduce the workload on teachers by providing tailored learning guidance.

In particular, edutech services like the "Integrated Learning Window", "Communication Collaboration Tool", and "Class Management Tool" are implemented, allowing teachers and students to share real-time content and facilitating various functions for interactive communication.

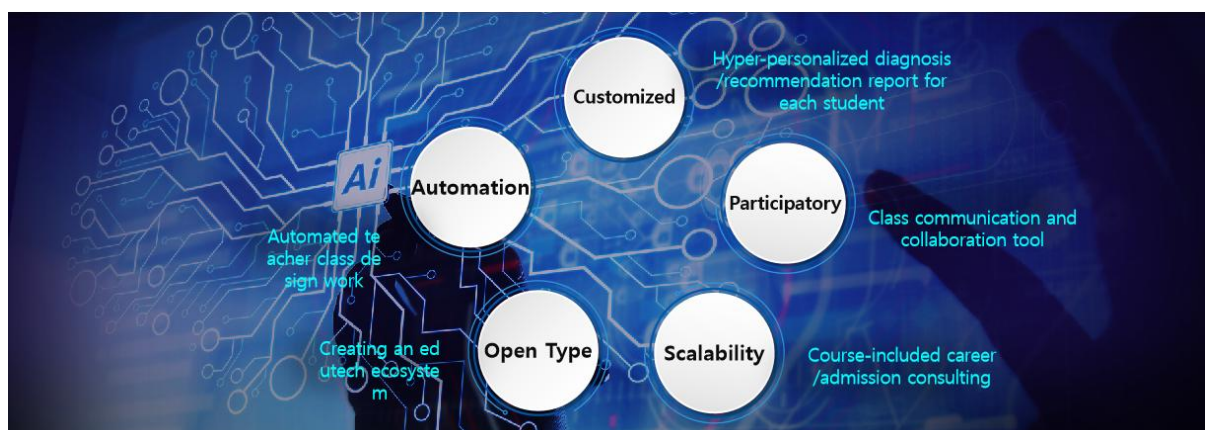


Figure 35 KT AI Future Education Platform

4.3.6 Health

As the era emphasizes the importance of individual health management in the face of an aging population and the increase in patients with chronic diseases, KT has developed a remote care service for chronic diseases and is currently pilot testing it. This service, staffed by professionally licensed nurses and nutritionists, sets achievable health goals for users. Care coordinators monitor the progress and encourage users, assisting them in managing their health and changing habits on their own. By analyzing various lifestyle data of the customer using AI, and in accordance with the care plans provided by medical professionals, the service provides health improvement goals and missions tailored to one's health condition. It also offers

additional services suitable for specific ailments, aiming to facilitate everyday health management effortlessly. Furthermore, KT has established a medical corporation in Vietnam, pushing forward into the global healthcare market by setting up health checkup centers and pursuing future business ventures.

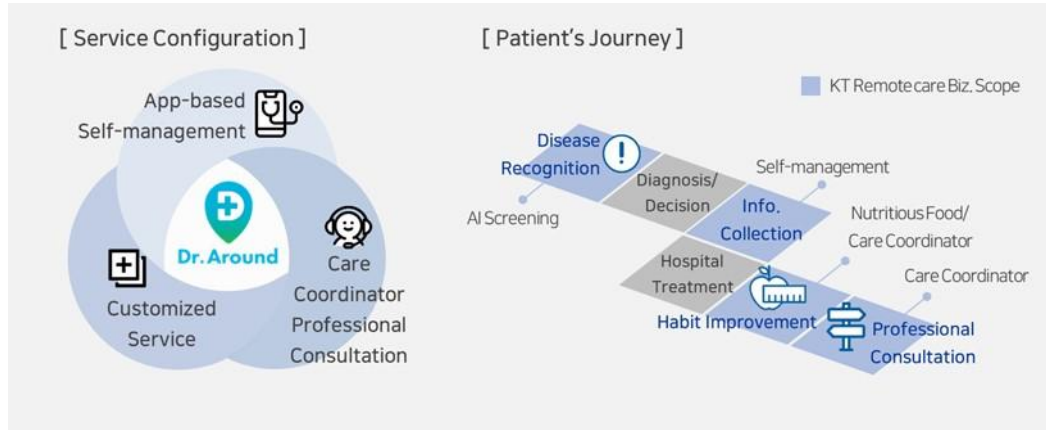


Figure 36 KT Remote Care Service

5. Industry Development Initiatives

Facing the future, China Mobile, NTT DOCOMO and KT look forward to collaborating with more industrial organizations and partners to jointly formulate long-term development vision, goals and principles, accelerate innovation in product, program and business pattern, promote industry consensus and build an open ecology. We launch the following initiatives:

1. AI technology and product development should follow the principles of fairness, non-discrimination, transparency and interpretability. On the one hand, human dignity must be guaranteed to make sure no discrimination or bias in data acquisition, algorithm design, technology and product development. On the other hand, the operating principles and decision-making standards must be articulated to provide understanding to users.

2. Resource and technology sharing should be enhanced through standardization and open ecology. More upstream and downstream industry partners are encouraged to participate in application innovation and help achieve technology breakthroughs and monetization. Open API and evaluation & certification mechanisms are promoted to lower the threshold of single-point innovation and system integration verification, and realized high-quality development.

3. Social responsibility should be reinforced to develop responsible artificial intelligence. The development of AI should aim at enhancing the common well-being of mankind, promoting the fair sharing of the benefits brought by technology development, and enhancing social fairness, justice and equal opportunities. AI Enterprises should also carry out continuous education and promotional activities on the ethical use of AI, enabling a sustainable development of the industry.

6. Abbreviations and Acronyms

AaaS	Ability as a Service
AGI	Artificial General Intelligence
AI	Artificial Intelligence
AICC	AI contact center
AIGC	AI Generated Content
API	Application Programming Interface
CSP	Communication Service Provider
IaaS	Infrastructure as a Service
LLM	Large Language Model
MaaS ¹	Model as a Service
MaaS ²	Mobility as a Service
MEC	Mobile Edge Computing
PaaS	Platform as a Service
RLHF	Reinforcement learning from Human Feedback
SDO	Standardization Organizations
SOE	State-Owned Enterprise