# White Paper of

# AISWare PEC V1.0

AISWare Privacy-Enhancing Computation PEC (hereinafter as AISWare PEC or PEC) enables enterprises to build trusted data circulation and transactions with data element value activation and bonus release

# Disclaimer Statement

AsiaInfo Technologies (China), Inc., hereinafter as "AsiaInfo Technologies" or “AsiaInfo”, exclusively owns all intellectual property rights, including but not limited to copyrights, trademarks, and patents, as well as technical secrets related to the Product and its derivatives, along with all related documentation, including all information within this document and any attachments.

The information within this document is confidential and intended solely for use by the recipient(s) designated by the user. Without prior written permission from AsiaInfo Technologies, any user of this document shall not take any actions with respect to this Product or the information contained herein to any third party, including but not limited to managers, employees, and affiliates other than the designated recipients. The aforementioned actions encompass but are not limited to developing, updating, compiling and decompiling, assembling, lending, transferring, selling, disclosing, authorizing, distributing, or any other actions. Nor shall any such third party be permitted to use the Product and the information in this document for any purpose whatsoever.

Without prior written permission from AsiaInfo Technologies, users shall not copy, modify, or distribute this document for any purpose. Altering, removing, or damaging any trademarks used in this document is strictly prohibited.

This document is provided as original, and AsiaInfo Technologies makes no warranties regarding the correctness, accuracy, reliability, or any other aspect of this document or its consequences after use. All information in this document is subject to further modification without prior notice, and AsiaInfo Technologies disclaims any responsibility for errors or inaccuracies that may be present in this document.

AsiaInfo Technologies shall not be held responsible for any and all types of liabilities, infringements, or damages resulting from the use of the Product or the information within this document. This exclusion of liability encompasses all forms of damages, including but not limited to direct, indirect, incidental, special, or punitive damages, regardless of whether AsiaInfo Technologies was notified of the possibility of such damages beforehand. The exclusion of liability applies to all forms of legal claims, including those arising from negligence or other torts.

AsiaInfo Technologies' products may include third-party software. Please refer to the copyright statements in the third-party software documentation for details.

## AsiaInfo Technologies Limited (Stock Code: 01675.HK)

AsiaInfo Technologies Limited ("AsiaInfo Tech") started in 1993 and was successfully listed on the Main Board of the Hong Kong Exchanges and Clearing Limited on December 19, 2018. As the largest provider of telecom software products and related services in China, AsiaInfo Tech has developed industry-leading R&D capabilities with a loyal customer base.

AsiaInfo Technologies (China) Inc., as an indirect wholly-owned subsidiary of AsiaInfo Tech, is a leading software product and service provider in China, boasting extensive experience in software product development and large-scale software engineering implementation. With 30 years of deep market presence, AsiaInfo has advanced technological capabilities and numerous successful cases in 5G, cloud computing, big data, artificial intelligence, the Internet of Things (IoT), smart operations, and business and network support systems. AsiaInfo's clientele spans across industries including telecommunications, broadcasting, energy, government, transportation, finance, and postal services.

In 2022, AsiaInfo acquired iResearch Consulting Group Co., Ltd. (iResearch Consulting) and integrated it into the new brand iDigital, expanding AsiaInfo's capabilities from product development, delivery services, data operations, and system integration to consulting planning and intelligent decision-making, establishing itself as a leading provider of end-to-end capabilities in digital intelligence.

AsiaInfo is committed to empowering various industries with technologies such as 5G, AI and big data, collaboratively creating digital value with customers. AsiaInfo aims to lead in both products and services, focusing on continuous product development in the areas of data and intelligence, cloud and network, IT, and middle office products. The cloud and network products maintain international leadership, while data and intelligence products achieve domestic leadership and some international advancements. In the IT domain, AsiaInfo's products stand at the forefront within the domestic landscape.

In the future, AsiaInfo strives to become the most trusted leader in digital intelligence, leveraging its comprehensive capabilities in the field to innovate customer value and contribute to the digital transformation.

## Certificates (Part)

Capability Maturity Model Integration (CMMI) Certificate Level 5 (L5)

Cloud Managed Services Capability Assessment Certificate: Excellent Level

Digital Trusted Services - R&D Digital Governance Capability Certificate

Enterprise Credit Grade (AAA) Certificate

Information System Construction and Service Capability Assessment CS L4

ISO9001 Quality Management System Certificate

ISO20000 IT Service Management System Certificate

ISO27001 Information Security Management System Certificate

Service Certificate of Information System Security Development L2

Service Certificate of Information System Security Integration L2

## Awards (Part)

Awards from International Telecommunication Union (ITU)

Award for Science and Technology Progress of Wu Wenjun Awards

Best Network Slicing Trail at 5G World Summit

French Design Awards

Global Telecoms Awards

IDC Future Operation Leadership

iF Design Golden Award of Hannover Industrial Design Forum

Leading Artificial Intelligence Enterprise in China

Leading Enterprise of Advanced Smart City

Outstanding Catalyst Contribution to TM Forum Assets

The Best Innovation and Future Techco of TM Forum

The Best Standard Contributor of TM Forum

The Most Innovative Application of AI & Automation of FutureNet Asia

The Most Influential Enterprise in China Software Industry

Top 100 China Software Business Revenue List for consecutive years

# Contents

# 1 Executive Summary

Privacy-preserving computation utilizes cryptography and distributed technology to protect data privacy while sharing computations. It is equipped with features such as FL, MPC, and trusted computation that enable deep data mining with privacy protection regulation compliance.

AsiaInfo Technologies has launched AISWare PEC, a privacy-preserving computation product with own IP and leading technologies to converge data securely in an " available and invisible " way under the technical trust mechanism. AISWare PEC adopts the pioneered privacy-preserving platform architecture “1+X” in the industry and provides functions such as FL modeling, PSI, PIR, and MPC.

Based on chips and trusted computation capabilities with the combination of self-developed software applications, AsiaInfo has innovatively developed a perfect and efficient end-to-end marketing solution for data elements for industries such as public security, new energy, finance, smart cities, telecom, and medical care, which effectively balances the need for data security and data sharing while protecting users' privacy and promoting sustainable data utilization and innovation.

AISWare PEC has been applied in diverse application scenarios, such as joint marketing, joint risk control, intelligent healthcare, and E-Government, and ensures secure and compliant data processing and analysis for enterprises and organizations, leading to increased business efficiency and innovation.

This White Paper will address the AISWare PEC regarding product overview, functional architecture, basic functions, featured functions, unique advantages, scenario solutions, and use cases.

# 2 Abbreviations and Terms

Common terms for *Product Name* are shown in Table 2-1.

**Table 2-1 Term explanation**

| Abbreviations and Terms | Full Names | Explanations |
|---|---|---|
| FL | Federated Learning | A distributed machine learning technique that enables model training among multiple local data sources without compromising data privacy. In this approach, global models are constructed through exchanging model parameters or intermediate results, except for local individual or sample data. This allows for the computation sharing while preserving data privacy, making it a new application paradigm of "data available but not visible" and "dynamic model except dynamic data". |
| MPC | Secure Multi-Party Computation | Participants use private data in confidential computation without disclosure to accomplish a particular computational task together. |
| PIR | Private Information Retrieval | A practical technique and application in secure multi-party computation that can be used to protect the user's query privacy and results, ensuring query completion on the premise while keeping the query information secure during user submissions to the data source. |
| PSI | Private Set Intersection | The two parties with data can compute the intersection portion of their data sets without disclosing any information beyond the intersection. |
| SQL | Structured Query Language | A computer language used to store, retrieve and modify data stored in a relational database. |
| TEE | Trusted Execution Environment | An environment that ensures the security, integrity, and tamper-resistance of code and data during the execution process. |

| Abbreviations and Terms | Full Names | Explanations |
|---|---|---|
| Blockchain | | Blockchain is a technological system maintained by multiple parties, using cryptography to secure transmission and access, and capable of consistent data storage, tamper-proofing, and anti-repudiation. |
| Interconnection network | | Network with cross-platform federated privacy computing services provided through interaction and collaboration between the connection of different privacy-preserving computation platforms after deployment. |
| Privacy-preserving computation | | Analysis and computation to guarantee the data "availability and invisibility" in circulation and convergence under the premise of ensuring that the data provider does not disclose the original data. |

# 3 Product Overview

AISWare PEC is a trusted data circulation platform that ensures availability and invisibility in controllable and measurable applications. Relying on technologies such as MPC, FL, and blockchain, it supports the creation of data circulation applications across diverse industries such as telecom, finance, government affairs, transportation, and energy. Through its advanced features, AISWare PEC enables the activation of data element values in these industries, thereby enhancing their efficacy and efficiency.

## 3.1 Trends and challenges

Currently, privacy-preserving computation technology is rapidly developing that can effectively address the data compliance issues faced by enterprises and organizations. It offers robust technical assistance for implementing data security systems. Nonetheless, it still grapples with the security, performance, and interconnectivity challenges of data. The development of privacy-preserving computation can be analyzed across three levels: technology, application, and regulation.

- Technology level

While there have been notable advancements in privacy-preserving computation technology, it has not yet been fully integrated into the industry. Many of these technologies remain in the laboratory or prototype stage and require further research and validation to improve their performance, scalability, compatibility, and usability. Additionally, the lack of unified standards and specifications presents challenges for interoperability and synergy between various technologies and platforms, leading to inefficiencies and lower-quality data circulation and analysis.

- Application level

In practice, it is not always straightforward to obtain sufficient quantities of data with high quality. Certain domains may face data scarcity, while certain datasets may have deficient values or incomplete information, which can negatively affect

the accuracy and robustness of machine learning algorithms and data-driven approaches.

- Regulation level

In regulation formulation relating to privacy-preserving computation, there is a need to balance the protection of individual privacy and the innovation and development of enterprises. Excessively stringent regulations may restrict the innovation and development of enterprises, while excessively loose regulations may make it difficult to protect personal privacy effectively. Therefore, targeted and operational regulations are necessary to facilitate the development and application of privacy-preserving computation technologies.

## 3.2 Product definition

AISWare PEC, based on the pioneered "1+X" architecture, can open and interconnect the heterogeneous operators and realize transparent and unified control of the operator components on the cloud-native architecture; it encapsulates complex cryptographic techniques in a unified way and provides a graphical development method for easier operation. Enterprises can leverage AISWare PEC to access valuable data assets and seamlessly integrate with production processes, and quickly build cross-industrial applications that prioritize privacy-preserving computation.

AISWare PEC delivers in two forms, AISWare PEC-Appliance, which is a combination of software and hardware, and AISWare PEC-Worker, which is a software platform. AISWare PEC-Appliance is compatible with servers that use x86 and ARM architecture and comes equipped with a variety of pluggable security enhancement modules, including hardware cryptographic module, TEE hardware module and trusted computing module, along with optional acceleration modules on-demand, such as GPU/FPGA/ASIC. AISWare PEC-Worker supports deployment methods in containerization, virtual machines, and so on.

## 3.3 Product positioning

AISWare PEC supports data computation from more than two parties. The data accesses PEC through the cooperation between the data 'provider' and privacy-preserving computation service 'vendor'; and The data 'vendor' provides calling services of privacy-preserving computation to the data 'user'.

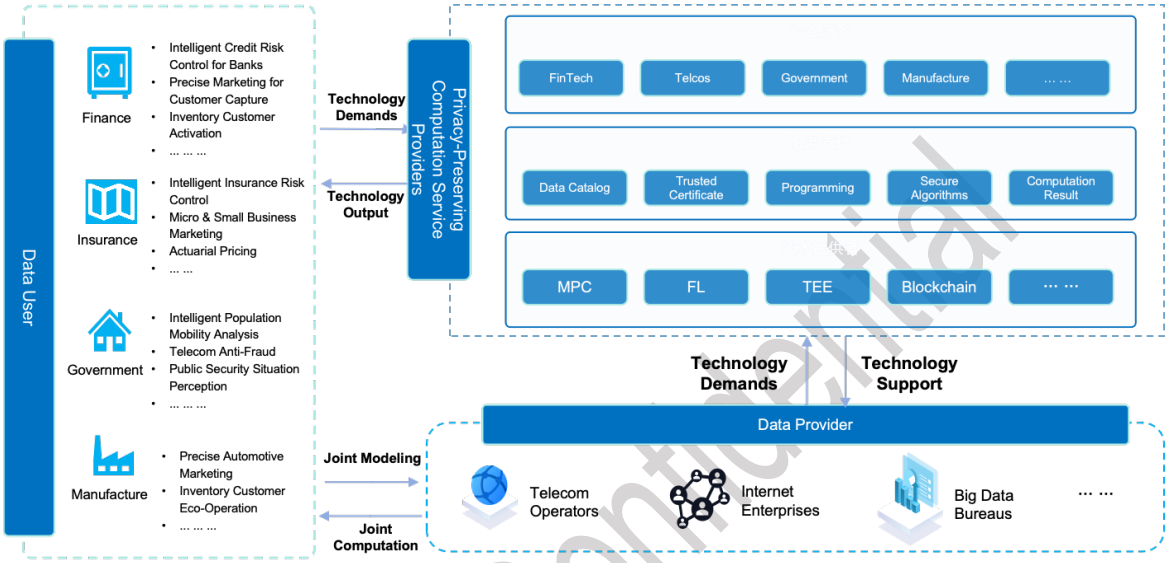



**Figure 3-1 AISWare PEC business diagram**

- Data provider: Enterprises with large amounts of data offer data through privacy-preserving computation products, such as telecom operators, Internet companies, and big data bureaus.
- Privacy-preserving computation vendor: Facilitate technology application through MPC, FL, TEE, and blockchain, and serve the finance, telcos, government, manufacturing and other industries.
- Data user: Different industries can engage in end-to-end privacy-preserving computation with data providers based on their business requirements.

# 4 Product Architecture

Based on data asset management, MPC, FL, blockchain, and other digital technologies, AISWare PEC enables seamless availability and invisibility of data, while also connecting data silos across enterprises and industries. By facilitating trusted data circulation and transactions, it enables value activation and unlocks the immense potential of data, thereby creating significant dividends and unleashing the energy of data elements.

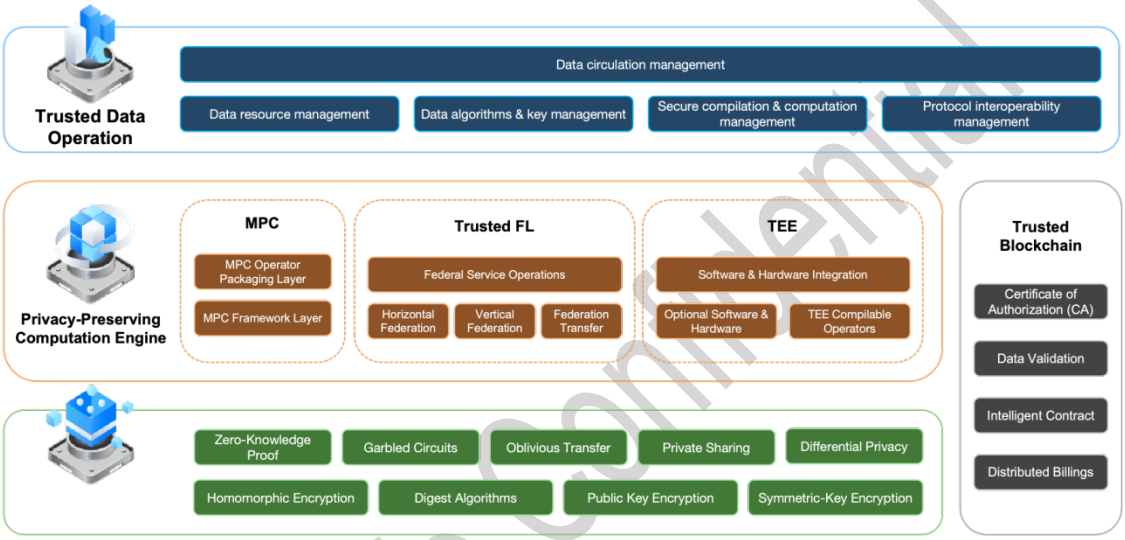



**Figure 4-1 AISWare PEC architecture**

- Primitive layer of communication cryptography: Encapsulation of cryptographic protocols based on homomorphic encryption, oblivious transfer, and confidential sharing in the primitive layer of communication cryptography.
- Privacy-preserving computation engine: MPC, trusted FL and TEE encapsulate core capabilities into applications.
- Trusted blockchain service: During the computation, the pre-computation and post-computation data are chained for data authorization and trusted data traceability.
- Trusted data operation: Include data resource management, security management, algorithm management, and protocol interoperability management, which satisfies the data security circulation.

# 5 Basic Functions

The basic functions of AISWare PEC include heterogeneous interconnection control, high-performance PSI, FL model development, operation service portal, hierarchical data control, and intelligent operation cockpit.

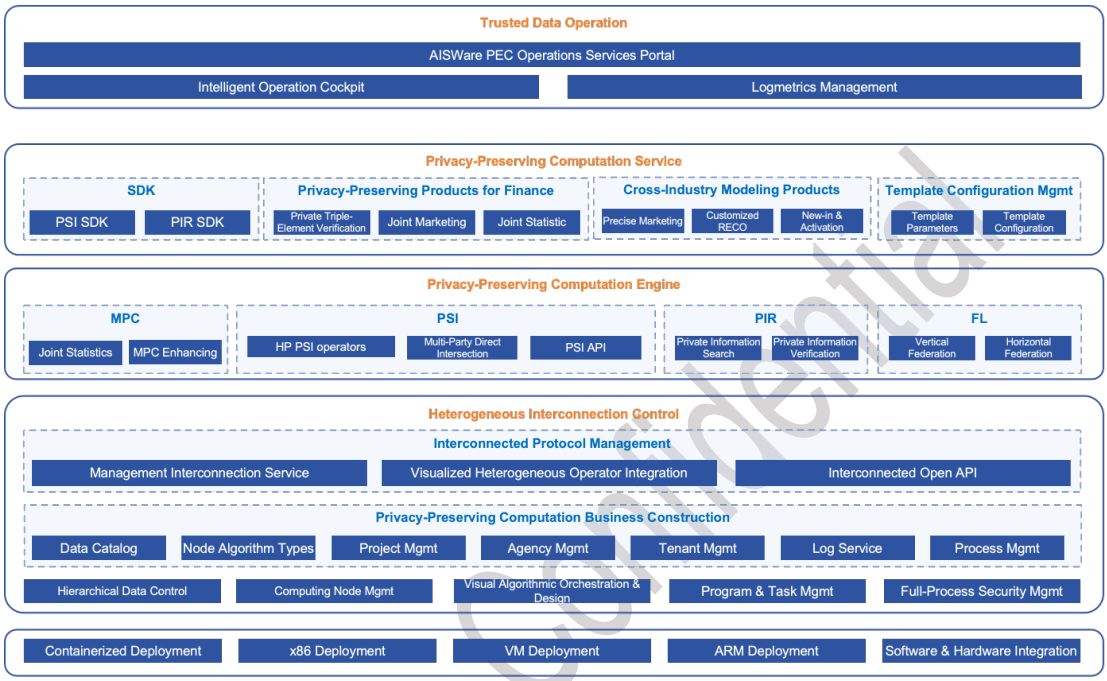



**Figure 5-1 Functional architecture of AISWare PEC-Worker**

- Heterogeneous interconnection control: Realize the business application of privacy-preserving computation products and the data circulation among various platforms based on the interconnection protocol management function and privacy computation construction function.
- High-performance PSI: Optimize computational resources, such as memory, to achieve highly-performed processing capability of PSI operators for billion-scale data computation.
- FL model development: Develop joint models based on homomorphic encryption with data not outbound for each participant, and support horizontal federation and vertical federation, as well as common machine learning algorithms such as logistic regression, decision tree, k-means, and neural network.

- Operation service portal: Apply for cooperation and entry according to the dataset information with support on entry progress inquiry and other functions.
- Hierarchical data control: Construct data security management systems, such as data categorization, as well as hierarchy and security strategies, to realize security governance and full life-cycle data protection.
- Intelligent operation cockpit: Analyze the general situation of cooperative operation and the operational situation of the two participant organizations with statistics on cooperative operation contents such as data, projects, operators and algorithms, and task monitoring.

As a specialized appliance combining software platform and hardware technology in one, PEC provides a hardware cryptography module, TEE module, and trusted computation module to support pluggable FL, MPC, TEE, and blockchain capabilities. Leveraging the telco's big data capabilities, it provides scenario-based service templates and out-of-box privacy-preserving computation services, which are ideal for empowering industries like finance, government, manufacturing, medical care, and telcos.

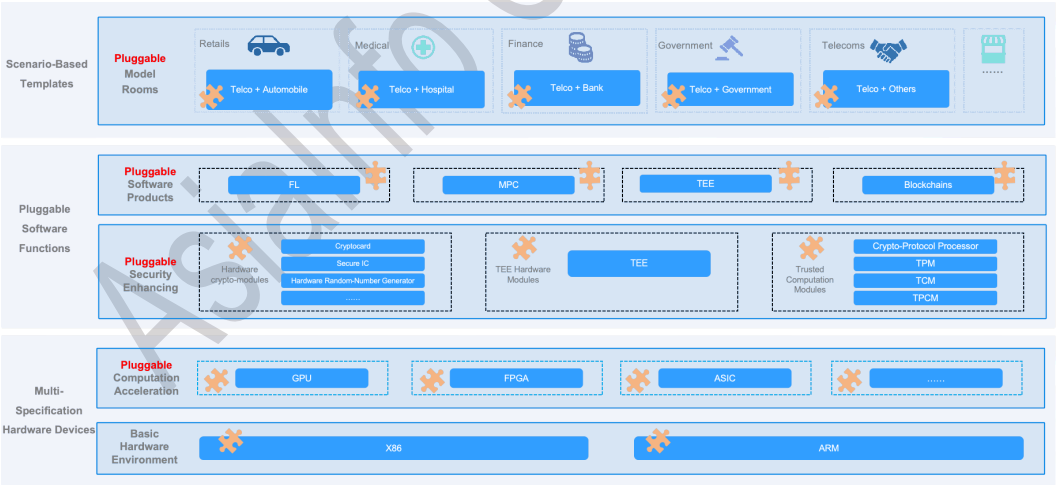



**Figure 5-2 Functional architecture of AISWare PEC-Appliance**

There are three specifications of AISWare PEC for different requirements, Mini, Standard, and Jumbo.


**Figure 5-3 AISWare PEC-Appliance specifications**

- Mini: Mainly target privacy data processing scenarios in small batches, with built-in MPC or FL, which can quickly provide customers with a basic privacy-preserving computation environment.
- Standard: Target at enterprise-level privacy-preserving computation scenarios, such as finance, government and retail, with MPC, FL, and blockchain capabilities, and providing out-of-box privacy-preserving computation services for enterprises.
- Jumbo: Support all enterprise-level privacy-preserving computation scenarios and provide professional security reinforcement and hardware co-acceleration with complex algorithms for satisfying the demands of large-scale data-volume privacy-preserving computation applications for the enterprise.

# 6 Featured Functions

The features of AISWare PEC include the pioneered “1+X” architecture for privacy-preserving computation, interconnected and hierarchical hosting, scenario-based privacy data service, scenario-based one-stop secure AI collaboration, software and hardware integration for collaborative acceleration, and chip adaption with self-research and reliability.

## 6.1 Scalable “1+X” Architecture

Pioneered “1+X” architecture for privacy-preserving computation platform integration (hereinafter as “1+X Architecture” or “the Architecture”) developed by AsiaInfo can interconnect and quickly integrate heterogeneous algorithms, and can quickly assemble cross-industrial privacy-preserving computation applications through visual orchestration interface. The Architecture consists of three parts: technical base, core functions, and open capabilities.

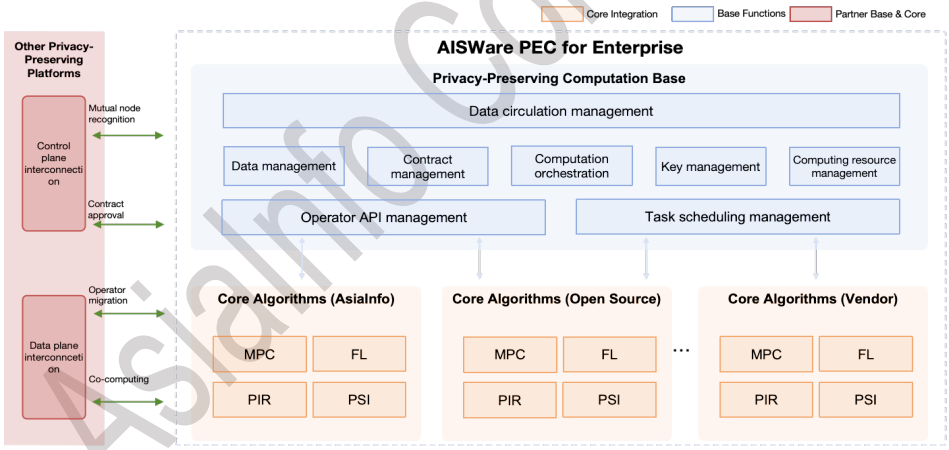



**Figure 6-1 “1+X” Architecture for privacy-preserving computation platform integration**

- The technical base is responsible for the management of data and its circulation, contracts, keys, operator API, resource and computation task scheduling.
- Core functions are responsible for the core algorithmic capabilities of MPC, PSI, PIR, and FL.
- The main scenarios, such as MPC, PIR, PSI, and FL, are opened by opening up the basic operator capabilities in the form of RESTAPI/RPC;

the technical base encapsulates the basic operator capabilities to form the external scenarios through the algorithm task orchestration.

In application, the Architecture integrates the core algorithmic functions of most of the mainstream privacy-preserving computation enterprises in the industry, which can achieve global data intelligence while maintaining the heterogeneous autonomy of each platform.

## 6.2 Interconnected and hierarchical hosting

AISWare PEC promotes the interconnection standardization process of privacy-preserving computation from the business value level. And it can utilize interconnectivity from the technical level with data privacy security.

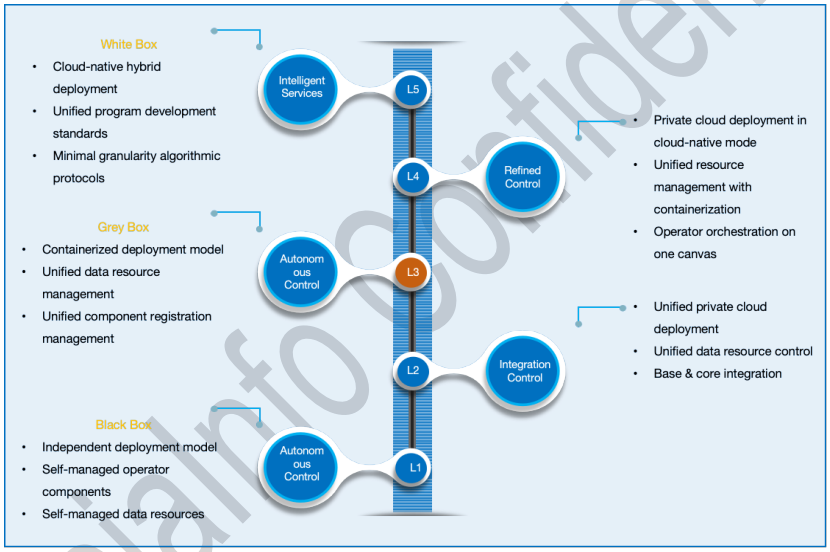



**Figure 6-2 Interconnected and hierarchical hosting**

Through the hierarchical hosting from L1 to L5, the interconnection mode between data resources, operator components and algorithmic protocols of heterogeneous privacy-preserving computation products transforms from black box to grey box, and from grey box to white box, which enhances multilateral trust and improves the efficiency of privacy-preserving computation interoperability in various industries.

## 6.3 Scenario-based privacy data service

By employing MPC, FL and other technologies, AISWare PEC can share and compute data with privacy security to meet the different demands in the financial industry; it can be applied to multiple scenarios, such as joint risk control, joint marketing, and anti-fraud.


**Figure 6-3 Scenario-based service**

## 6.4 Scenario-based one-stop secure AI collaboration

AISWare PEC provides the full-process pull-through capability of trusted FL from data preparation, federation establishment, and federation training to model deployment and federation reasoning, which facilitates the low-threshold implementation of federation modeling capability in vertical industries through open and inclusive federation model development, application and service.

- Visualized federation model orchestration: Provide drag-and-drop federation model development function, which allows developers to complete federation training task development through simple interface operation and configuration without complex configuration file editions.
- Open and inclusive: Compliantly access to multiple storage types and multiple data scales; support heterogeneous computation engine interconnection and interoperability, docking and expansion of algorithm components of heterogeneous platforms, and hot-pluggable

configuration of algorithm components; allow enterprises smoothly access their data in the production environment with significant data docking cost reduction.

## 6.5 All-in-one system for synergy

AISWare PEC adopts synergetic computation in hardware and software integration, with significant speed performance improvement in model training and cryptographic computation, along with 5 to 10 times in algorithm performance improvement. It also supports co-acceleration of heterogeneous computing power and transfers complex calculations to hardware devices for execution, which significantly enhances the processing efficiency of algorithms in parallel with high concurrency and low latency.

- Plug-and-play heterogeneous hardware acceleration: Support PPU co-acceleration such as FPGA, DCU and ASIC.
- Significant end-to-end performance improvement: 5 to 10 times improvement in overall end-to-end performance and more than 10 times improvement in individual arithmetic performance with external hardware PPUs.
- Flexible and user-friendly application layer development: Support various homomorphic computing operators, and construct Numpy heterogeneous acceleration operator interfaces to support module API calls.
- Software and hardware synergy to accelerate multiple scenarios: Support various privacy-preserving computation tasks such as privacy query, PSI, feature engineering, joint modeling, and joint prediction.

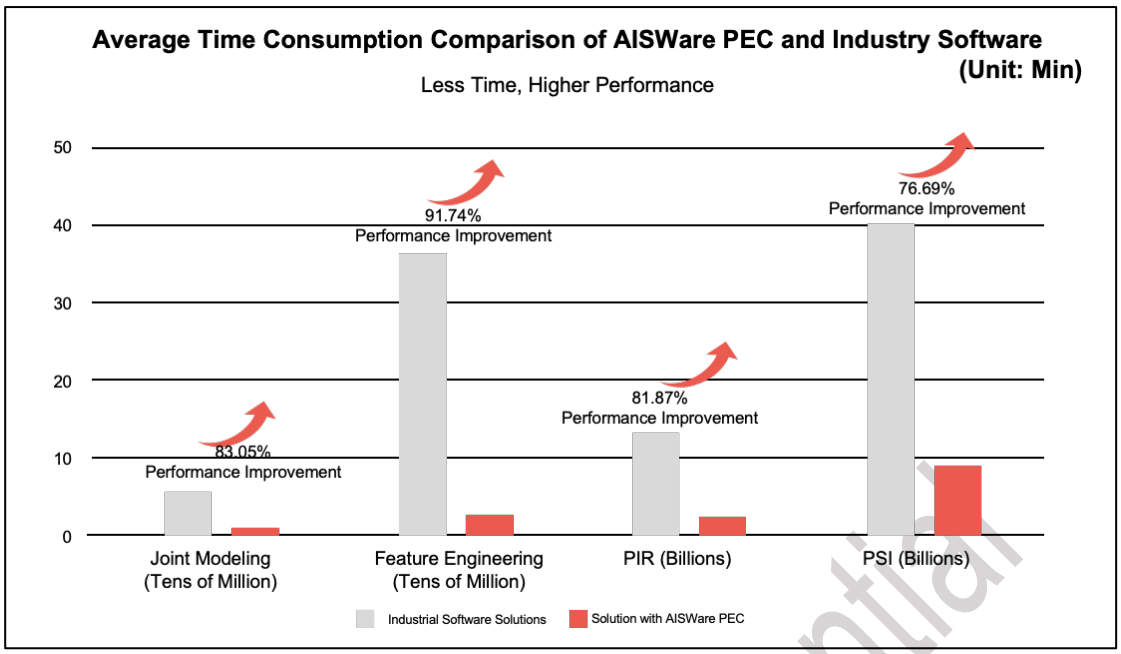



**Figure 6-4 AISWare PEC performance comparison**

## 6.6 Adaptive algorithm on heterogeneous hardware

AISWare PEC supports hardware chips, operation systems and middleware in full-stack and full-process coverage with high performance, reliability and security. It also independently researches and develops privacy-preserving computation algorithm software and solutions such as MPC, PSI, PIR, and FL based on cryptography primitives.

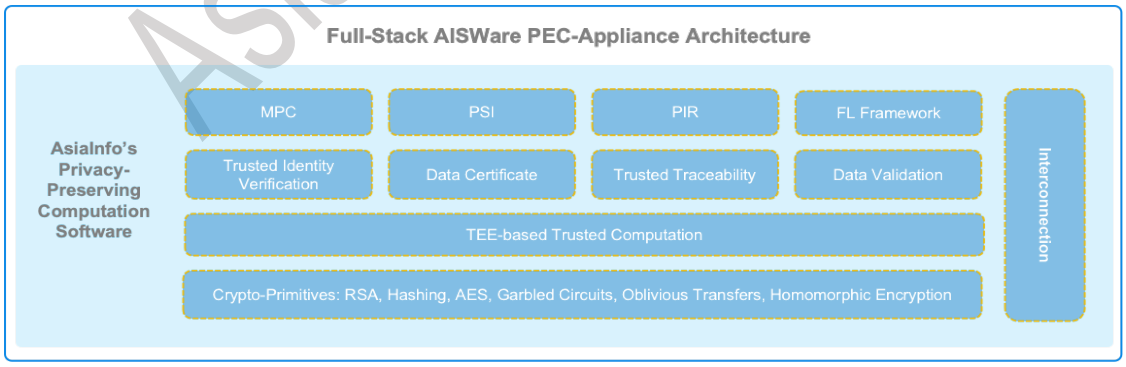



**Figure 6-5 Full-stack technical system of AISWare PEC**

# 7 Unique Advantages

The professional advantages of AISWare PEC in privacy-preserving computation can be seen in the openness and interconnection, standard leadership, out-of-box, and encryption in high performance.

## 7.1 Openness and interconnection

As an innovative solution provided by AISWare PEC, the 1+X Architecture supports pluggable one-click integration of multiple industrial algorithms, which provides a flexible, efficient, and secure platform in the privacy-preserving computation field. Through this Architecture, users can easily integrate operators from different vendors to realize cross-platform and cross-device privacy-preserving computation.

- Flexibility: 1+X Architecture allows flexible selection and combination of different algorithms and components according to different needs and scenarios under the premise of guaranteeing core functionality, which can greatly improve the adaptability and flexibility of the platform.
- High efficiency: 1+X Architecture maximizes computational and communication efficiency by optimizing the interaction between algorithms and components for more efficient data processing and analysis.
- Openness: 1+X Architecture supports extensive openness and extensibility for easy integration and interaction with other platforms, systems and applications, thus fully utilizing existing technologies and resources and reducing development and integration costs.

## 7.2 Industry technical standard-driven

1+X Architecture created by AISWare PEC provides a unified standard (hereinafter as “the Standard”) for interconnection and interoperability in the privacy-preserving computation field to realize data sharing and computation

between different platforms for better cross-industrial data cooperation and innovative development.

The Standard integrates the core algorithmic functions of most mainstream privacy-preserving computation enterprises to achieve global data intelligence while maintaining the heterogeneous autonomy of each platform. This hierarchical hosting mode supports the loose coupling design of management systems, algorithm protocols and computational primitives, which in turn enables the security visualization of hierarchical interconnections, improves the interpretability at the security level, and allows users to master better system operation capabilities.

In addition, the Standard further refines the interface specification system. It forms the cross-platform interconnected and hierarchical hosting mode from L1 to L5 under 1+X Architecture, and each level has corresponding requirements and provisions to meet the interconnection needs under different scenarios.

## 7.3 Out-of-box demos

AISWare PEC is pre-installed with industrial model rooms for multiple domains, including privacy-preserving computation solutions for different industries and different data types and applications, such as enterprise risk control, precise marketing, and customer attrition; it can provide an excellent experience with one-click deployment and out-of-box.

## 7.4 High-performance encryption

AISWare PEC utilizes various efficient encryption techniques to ensure data security and privacy, including homomorphic encryption, zero-knowledge proof, and MPC computation, which enable data operations such as encryption, decryption, computation, and verification, while ensuring the accuracy and security of the computation results.

- Homomorphic encryption is a cryptographic technique capable of additive and multiplicative operations with computing result accuracy except for plaintext data exposure.

- Zero-knowledge proof is a cryptographic technique capable of authenticity and validity verification of certain data or information without any exposure of the data or information.
- MPC is a cryptographic technique capable of joint computation without data exposure from multiple participants, allowing the authenticity and security of the computation results.

Highly efficient encryption technology, as one of the key features of AISWare PEC, enables comprehensive data protection and secure computation for data privacy-preserving and security.

## 7.5 Quick replication of application scenarios

AISWare PEC provides quick replication capabilities for cross-industrial collaboration. Based on the business accumulation and data comprehension in the telecom field as a production-grade tool, it expands cross-domain collaboration applications horizontally through out-of-box scenario templates and flexible and lightweight deployment capabilities.

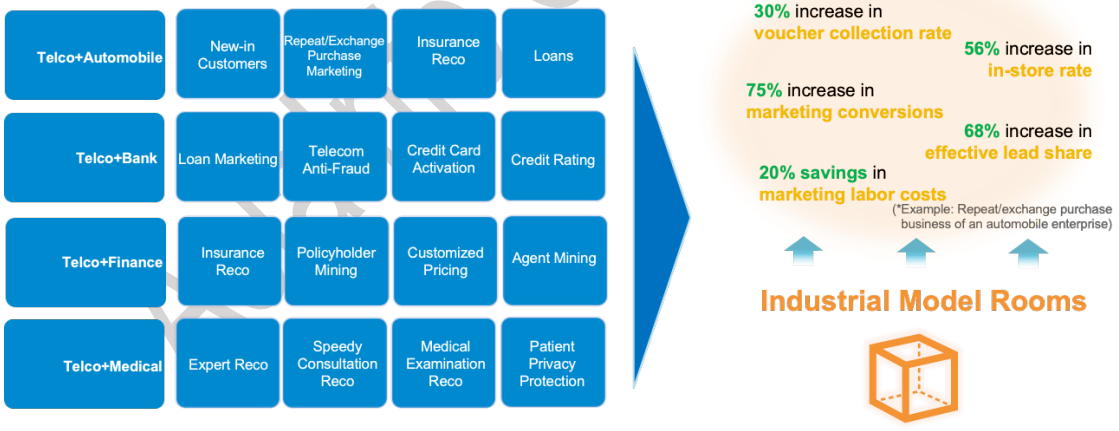



**Figure 7-1 Quickly replicable industrial demo rooms**

# 8 Scenario Solutions

Solutions with AISWare PEC provide rich application scenarios for telcos, finance, government affairs, energy, and automobile enterprises, including several major categories of application scenarios, such as PIR, PSI, joint statistics, and joint modeling, which empower the enterprise transformation of data intelligence through refined management, lean production, precise marketing, and precise planning.

The scenario solutions will be introduced in three parts, application scenario, business requirements and solution.

## 8.1 Private Information Retrieval

Private Information Retrieval (hereinafter as PIR), as a sub-branch under MPC in the field of privacy-preserving computation, makes it impossible for data holders to know the specific query object to better protect the privacy information of the querying party, which eliminates security risks and promotes the secured and organized data circulation.

### 8.1.1 Application scenario

PIR is mainly applicable to label queries, rating queries, list queries, information verification, and other application scenarios.

### 8.1.2 Business requirements

By introducing secure computation, during the PIR process, the querying party can only get a query result from the data service provider, and the data service provider cannot trace the query itself, which can effectively protect the privacy of both parties.

- Anonymity maintenance: Customers or users are not required to provide their real identification information during the query process, but rather the query is conducted anonymously to protect their privacy and security.

- Query efficiency improvement: Under privacy protection and security assurance, query efficiency improvement is another crucial business requirement.
- Data security assurance: During the query process, data security and integrity need to be ensured to prevent data from being tampered with or exposed.
- Accurate query results: Despite anonymous querying, the results should accurately reflect the customer's actual situation and needs for better customer service.
- Real-time query: Customers or users require real-time query to the latest data information to make timely decisions and responses.

### 8.1.3 Solution

PIR is a method for data privacy protection with the core idea of data query and utilization without disclosure of query requirements and identification information.

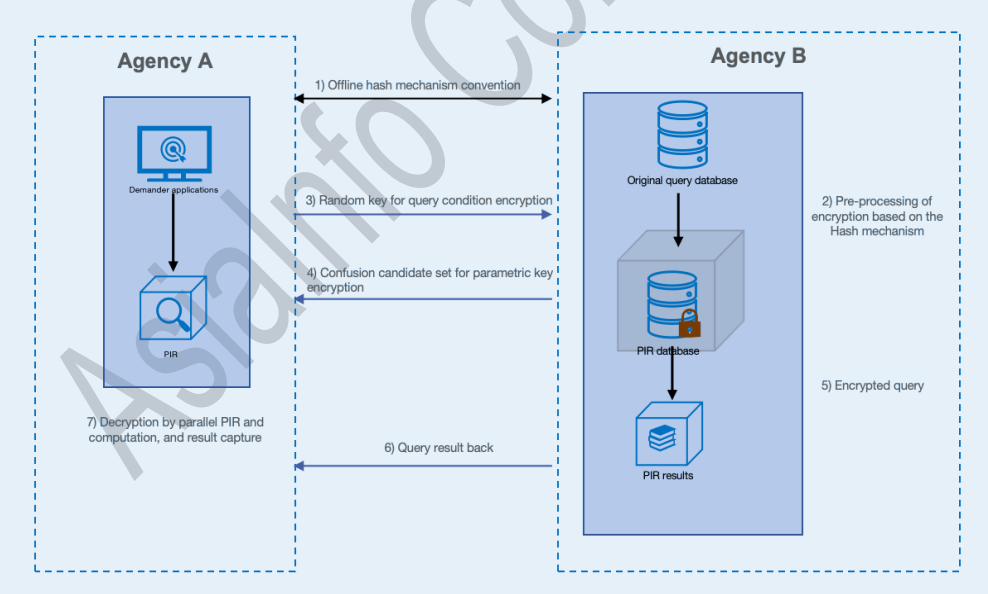



**Figure 8-1 PIR solution**

- Generating base key by the querying party: The querying party generates the base key based on the public key of the data holder, its public key, and the serial number of the queried data.

- Sending base key to data holder: The querying party sends a basic key to the data holders to make them determine the recovery keys for the data serial numbers in different retrievable data based on the inverse of a preset derivation algorithm.
- Determining the recovery key by the data holder: The data holder determines the recovery key based on the inverse of the preset derivation algorithm and the base key, then determines the encryption key for the different retrievable data.
- Encrypting data by the data holder: The data holder obtains the ciphertext of each retrievable data by encrypting the corresponding retrievable data with each encryption key.
- Obtaining each retrievable data ciphertext returned by the data holder: The querying party obtains each retrievable data ciphertext returned by the data holder.
- Parsing the ciphertext by the querying party: The querying party parses the query according to the holders' public key and recovery key.

## 8.2 Private Set Intersection

In Private Set Intersection (hereinafter as PSI) scenario, each participant performs an intersection process on their data. The intersection computation process only protects the information privacy out the intersection, meaning that each participant will get the same part of the other's datasets as theirs except for the different part.

### 8.2.1 Application scenario

The application scenarios mainly include target user alignment, PSI, secure union computation, labeled data expansion, and so on.

### 8.2.2 Business requirements

The business requirements of PSI for privacy-preserving computation include data privacy protection, data control preservation, intersection computation

accuracy, computation efficiency, security and compliance, which need to be considered and satisfied comprehensively when applying the PSI technique.

- Data privacy protection: During the PSI process, it is necessary to protect the privacy information of the raw data to avoid data exposure and misuse.
- Data control preservation: PSI for privacy-preserving computation should preserve the data control of each party, which means that each party can only access its data without exposure to others.
- Intersection computation accuracy: The PSI result should be accurate, meaning that the calculated intersection result should be correct without errors or omissions.
- Computation efficiency: PSI needs to keep the computation efficiency to meet the business requirements.

### 8.2.3 Solution

As a specific application in the field of privacy-preserving computation, PSI implementation is based on the intersection technique of privacy-preserving set that allows two or more parties holding their respective sets to jointly compute the intersection; at the end of the protocol interaction, one or more of the parties obtains the correct intersection according to a pre-agreement without any information outside of the intersection from other parties.

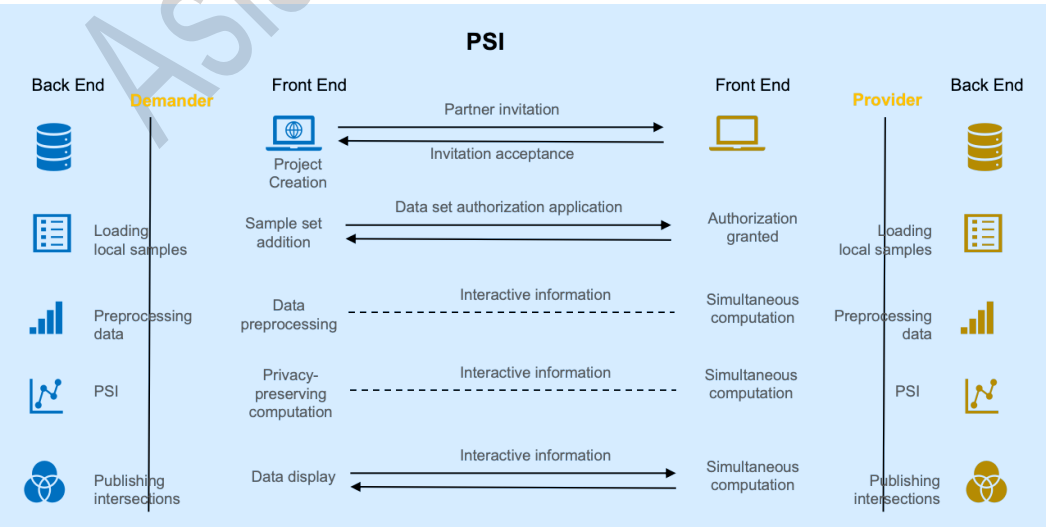



**Figure 8-2 PSI solution**

- Construct pseudo-random functions based on oblivious transfer extensions to compare data, and use the Cuckoo hash algorithm to reduce the volume of data transferred.
- Compared to the PSI implemented by the basic oblivious transfer algorithm, the performance significantly improves with a substantially reduced network overhead.

## 8.3 Joint statistics

Joint statistics can facilitate participants to conduct statistical calculations without exposing their respective data, thus obtaining more accurate results.

### 8.3.1 Application scenario

There is a wide range of application scenarios for joint statistics. For example, in the total financial asset authentication scenario, banks and securities can jointly conduct asset statistics to better assess the asset status and credit risks of their clients; in the medical statistics analysis scenario, hospitals and research institutes can jointly conduct statistical research on disease incidence and prevalence trends to better formulate prevention and treatment plans; in the resident population migration scenario, the government can jointly conduct demographic statistics and analysis to better manage urban development and the public resource allocation.

### 8.3.2 Business requirements

Joint statistics can prevent the raw data of all parties from being exposed. In the case of untrusted third parties, multiple participants are involved to accomplish collaborative computation securely, which means, multiple data parties perform statistics in a distributed network, and the user parties have no additional information about the statistics but statistic results(labels). It is applied in several industries:

- Financial industry: Joint statistics can be used for risk assessment, investment analysis and market forecasting. For example, banks and securities can use joint statistics to calculate their clients' total assets

and return on investment to better assess their credit risk and investment value.

- Medical industry: Joint statistics can be used in statistical studies of disease incidence and prevalence trends, as well as in the utilization and allocation of medical resources. For example, hospitals and research institutes can use joint statistical studies on the incidence and transmission of influenza in a particular area to better develop prevention and treatment plans.
- Governmental affairs: Joint statistics can be used in urban planning, public resource allocation and population migration. For example, governments can use joint statistics to calculate the number and distribution of resident population in various city areas for better planning of public facilities and resource allocation.
- Marketing: Joint statistics can be used in consumer behavior analysis, market research and advertising effectiveness evaluation. For example, e-commerce platforms can use joint statistics to calculate consumers' purchasing behaviors and preferences to better recommend products and optimize advertising.
- Social media: Joint statistics can be used for user behavior analysis, topic trend research and social relationship analysis. For example, social media platforms can use joint statistics to calculate topic attention and activeness of users to better recommend content and optimize user experience.

### 8.3.3 Solution

Based on the requirements, data are collected from various providers and pre-processed for joint statistics. For data privacy protection, the noise addition and encryption of the data are required by using some privacy protection techniques such as differential privacy and homomorphic encryption. Secure data transmission and sharing is achieved through techniques such as secure channels and encrypted transmission to ensure data security and privacy

protection. Joint statistics can be implemented through appropriate statistical methods and algorithms, such as joint distribution and joint mean.

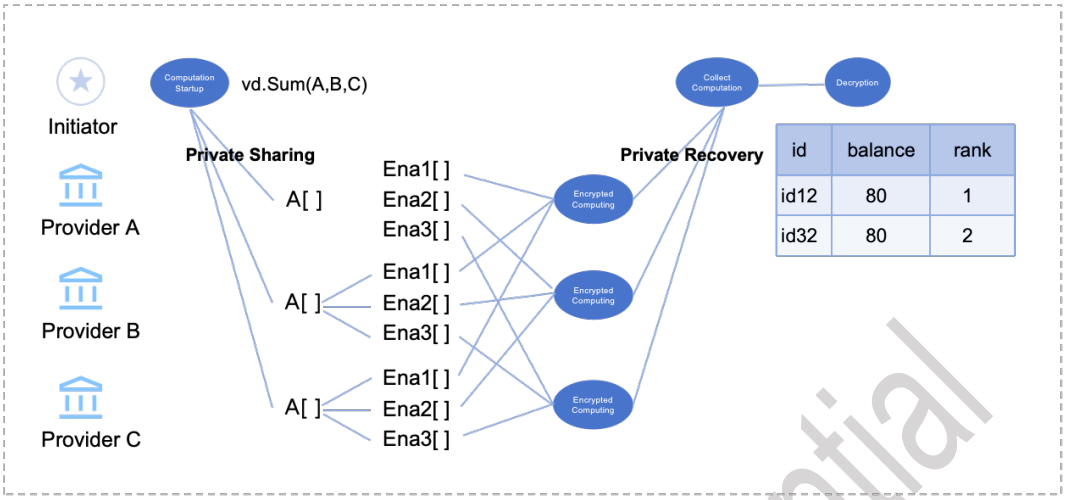



**Figure 8-3 Joint statistics solution**

Implementation procedure: Offline agreement on statistical analysis algorithms; Separately process after negotiation of pre-set numbers; issuance of pre-set numbers; model grouping computation and random splitting of results; sharing portions in secret; result portions; negotiation of secret portions and recovery.

## 8.4 Telco+Automobile: Joint marketing

After separate access to the dataset by automobile enterprise and telco, modeling is completed through secure data alignment and vertical FL, and the telco's feature contribution can be assessed through key indicators (such as IV value); after releasing the model, the repeat/exchange purchase marketing system makes a service call through the certified API to improve the marketing accuracy through joint modeling.

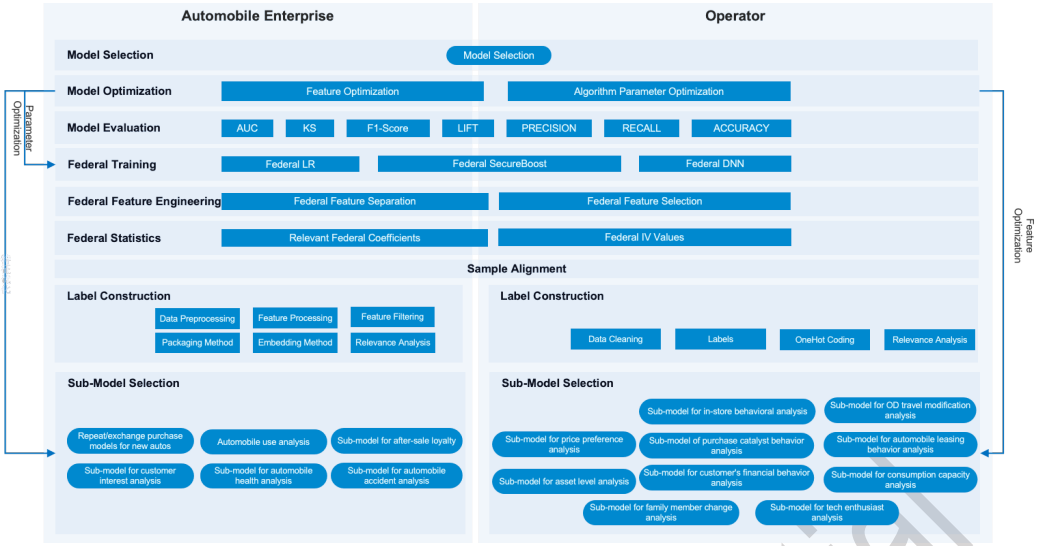



**Figure 8-4 Repeat/exchange Purchase Marketing Solution for Intelligent Automobile**

## 8.5 Telco+Bank: Credit checking

Under the premise of secure data privacy, the telco and bank both achieve vertical joint modeling of potential loan user models without data sharing, technically breaking data silos and achieving AI collaboration.

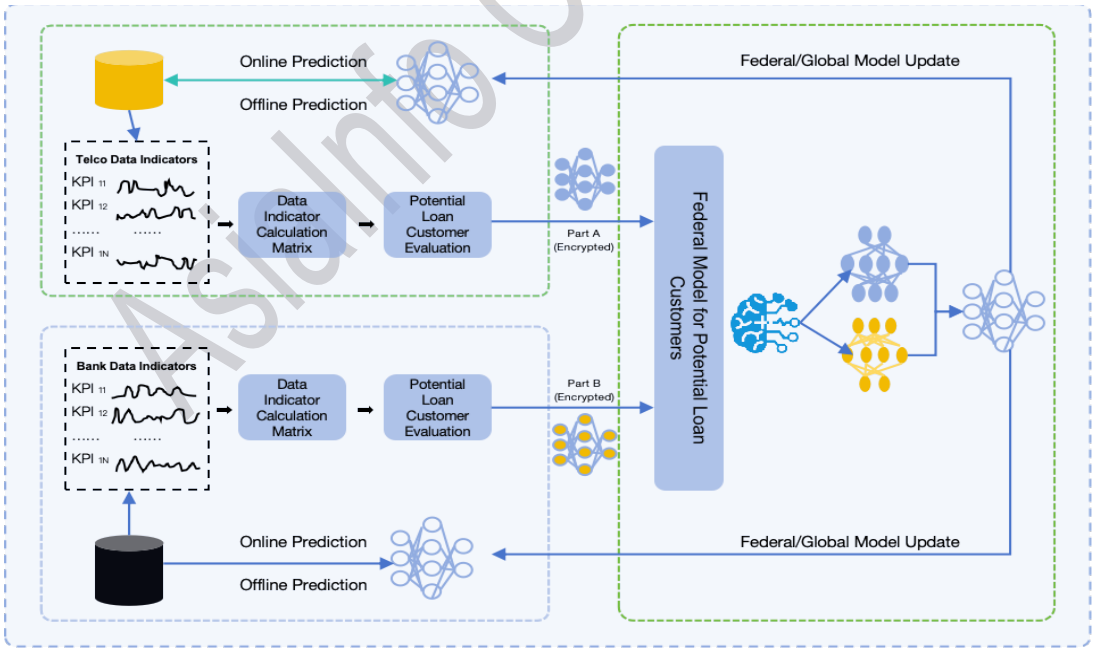



**Figure 8-5 Telco+Bank solution for credit rating scenarios**

FL modeling far exceeds single-domain modeling, approaching the modeling's theoretical optimal solution performance. In terms of accuracy checking, FL

modeling improves by 10% compared to a single bank domain and 30% compared to a single telco domain; and in terms of completeness checking, FL modeling improves by 5% compared to a single bank domain and 10% compared to a single telco domain.

## 8.6 Telco+Telco: Intelligent Anti-Fraud

The main reasons for telecom fraud detection difficulty are, on the one hand, the relative scarcity of fraud data, and on the more critical side, the lack of data interoperability between telcos, which forms data silos. Under the premise of no disclosure of data privacy (128-bit security) with FL, the user data from telcos are utilized to extract the fraud features, and the classification model is jointly trained based on the horizontal FL algorithm.

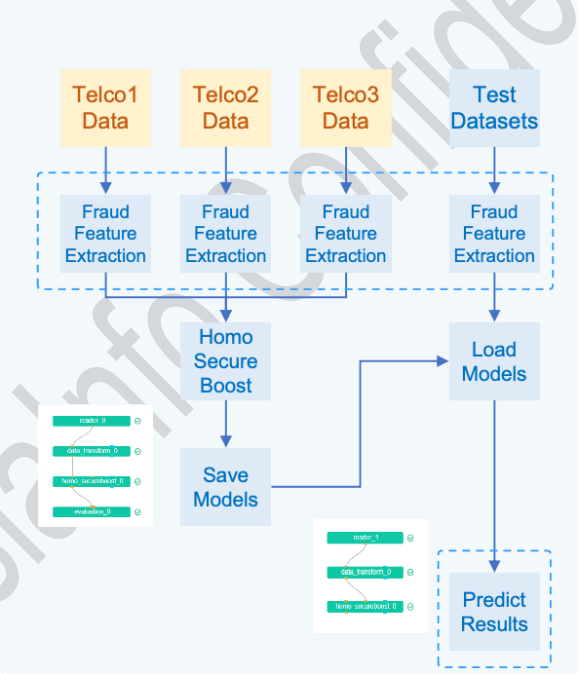



**Figure 8-6 Intelligent Anti-Fraud solution**

## 8.7 Telco+Insurance: Insurance agent mining

Aiming at solutions to low retention rates and high turnover of newcomers caused by the low business quality of insurance enterprises, the data of both telcos and insurance enterprises are used to mine out the most suitable group of prospective newcomers for insurance marketing to enhance income.

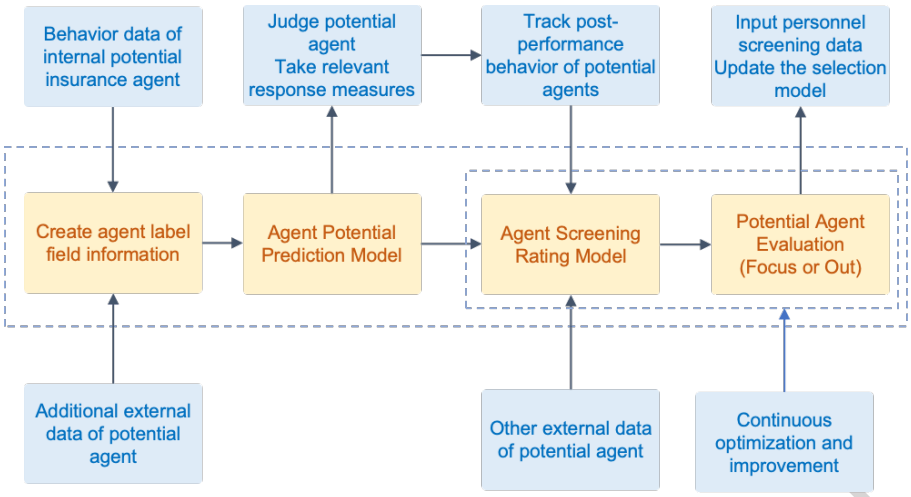



**Figure 8-7 Insurance agent mining solution**

## 8.8 Telco+Medical: Intelligent Recommendation

Based on the FL model architecture, cross-domain modeling capabilities are built for telcos and medical institutions with data privacy security assurance, which empowers intelligent medical recommendation scenarios by customized recommendations to different users of expert consulting, speedy consultation, and medical examinations.

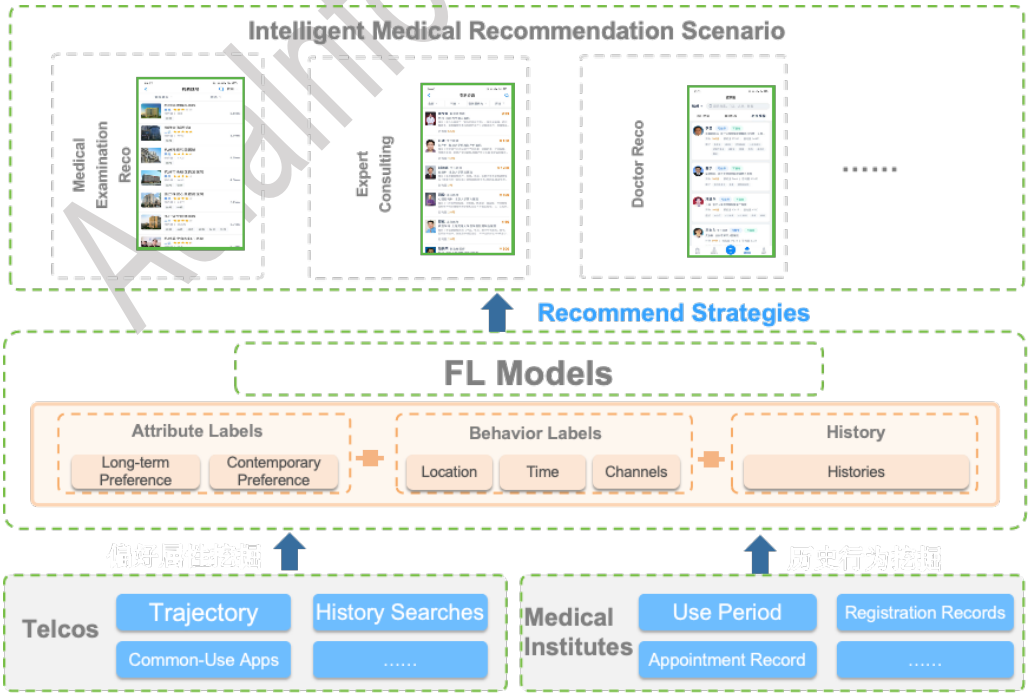



**Figure 8-8 Intelligent Medical Recommendation**

# 9 Use Cases

AISWare PEC can be applied to several industries, such as telecom, finance, government, medical, and manufacturing.

## 9.1 Secured data delivery platform for telco

An operator has accumulated valuable data assets with the features of ultra-full coverage, ultra-multi dimensions, uninterruptedness, and mega-scale while facilitating economic and social transformation. At the telecom operator level, it provides the foundation for the interconnection of privacy-preserving computation platforms to form an overall external data service capability.

### 9.1.1 Client requirements

The privacy-preserving computation platform (hereinafter as “the Platform”) construction for the operator is dual-driven by strategy and business. On the one hand, data circulation and data element marketing are encouraged nationally, and data security regulation is becoming increasingly stringent; on the other hand, operators have a natural data advantage, and vertical industries urgently need data to conduct in-depth analysis for industry data convergence and decision-making.

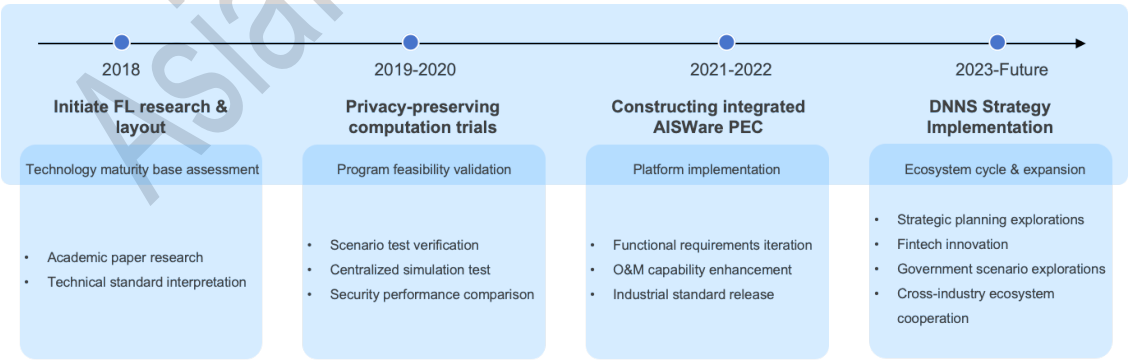



**Figure 9-1 Project background**

Though there exist strong demands for data convergence, data attribution is fragmented, preventing data from value maximization; and there are difficulties

in data circulation due to data privacy, data security, data performance, security compliance, and so on.

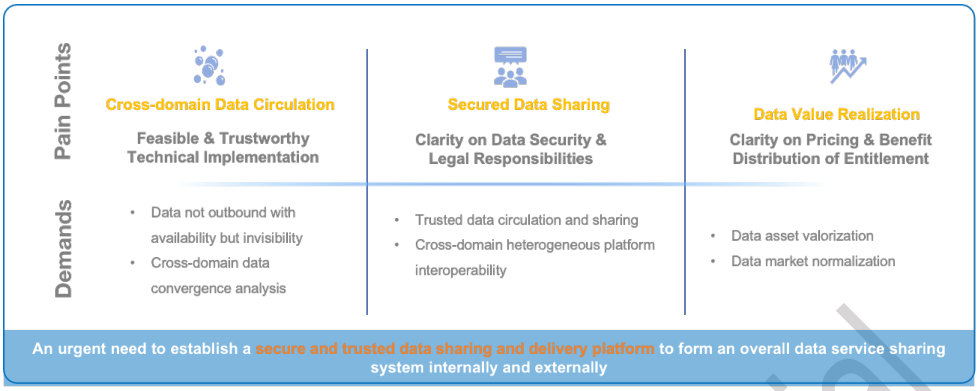



**Figure 9-2 Pain points and requirements of the client**

## 9.1.2 Solution and effects

Based on privacy-preserving computation technology and the "Agile Concept", AsiaInfo has provided the cross-industrial data convergence service under "1+X" Architecture, which enables the operator to jointly innovate with various industries with a big data ecosystem and development.

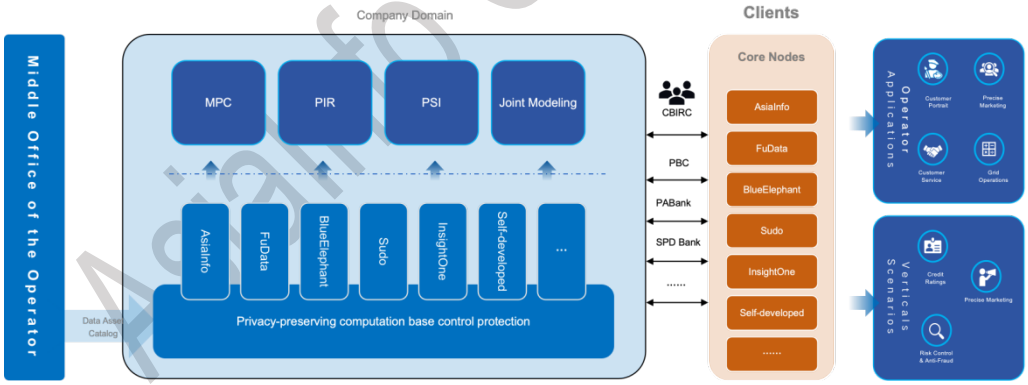



**Figure 9-3 Solution**

Initially completed in 2022, the Platform has served eight industries and supported 11 types of cross-industrial application scenarios; it has gradually developed and expanded a new ecosystem for openness and cooperation of China Mobile's data elements with substantial economic and social benefits.

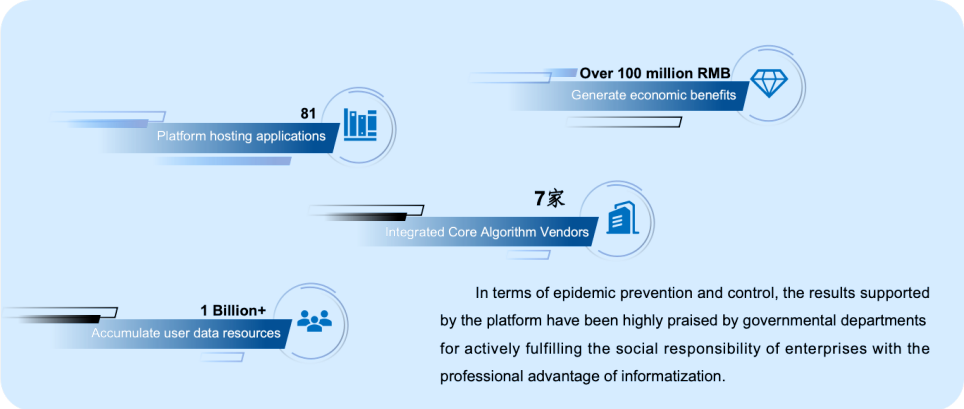



**Figure 9-4 Project effects**

## 9.2 Privacy-preserving computation for a financial institution

The whole business process of the financial institution, which consists of marketing, operation, and risk control, can converge its data with other institutions based on privacy-preserving computation technology to greatly enrich the financial scene and extend the financial industry.

### 9.2.1 Client requirements

The financial institution has accumulated a large amount of high-value data such as transaction, banking, market, risk, and customer data. However, in scenarios of marketing, operation and risk control, the data owned by the institution is relatively single and limited. To expand its business scenarios, there is a strong need for external data complying with regulatory requirements for data security.

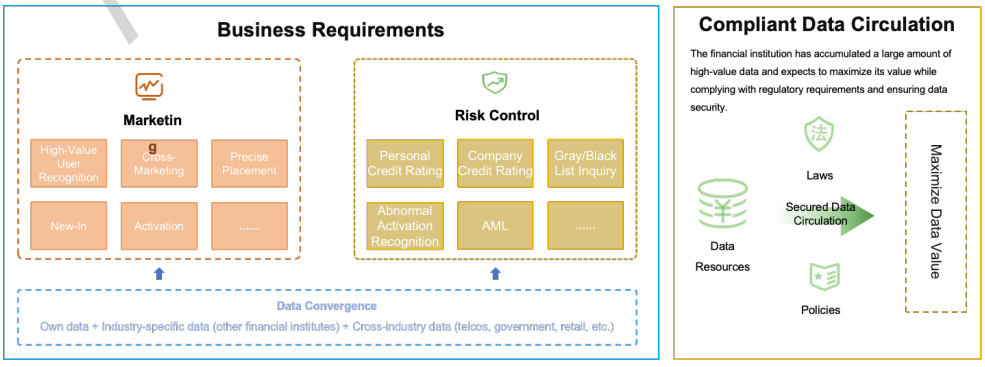



**Figure 9-5 Client requirements**

- In terms of marketing, the financial institution needs to understand customer demands and preferences to provide customized marketing services. Through AISWare PEC, it can unite external data for customer portrait analysis and risk assessment under data privacy-preserving to provide more accurate marketing services. Simultaneously, the financial institution can conduct joint marketing with external partners for data and resource sharing to improve marketing effectiveness and efficiency.
- In terms of risk control, the financial institution needs to assess and manage risks for loss prevention. Through AISWare PEC, it can conduct risk assessment and anti-fraud analysis to detect and handle risk events promptly under data privacy protection. Simultaneously, the financial institution can share and distribute risk with external partners to improve risk management capability and enhance efficiency.

### 9.2.2 Solution and effects

Relying on AISWare PEC, the financial institution can introduce horizontal (cross-industry) and vertical (intra-industry) institutions as partners to establish an open data ecosystem under its control. PEC adopts 1+X Architecture, which realizes platform openness and interoperability by integrating the algorithmic functions from other core data vendors.

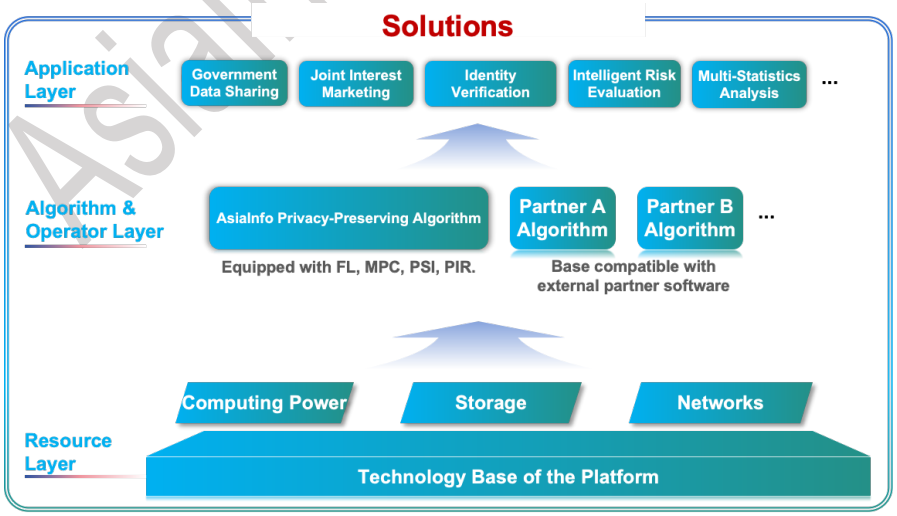



**Figure 9-6 Construction solution**

- Resource layer: Unified control of institutions, nodes, algorithms, projects, programs and tasks based on AISWare PEC to realize applications such as computation power and data storage.
- Operator and algorithm layer: Provide algorithms of PSI, PIR, joint statistics and FL developed by AsiaInfo, and simultaneously integrate algorithms from external vendors based on the "1+X" Architecture.
- Application layer: Apply base data in scenarios through algorithms and telcos, such as telecom anti-fraud, joint marketing, three-authentication verification, intelligent risk control, joint statistics and other scenarios.

The financial institution, as a data demand party, can enhance the marketing risk control effect by enriching user-profiles and gray and black lists; as a data provider, it can open the data to other members in the ecosystem with corresponding revenues to maximize its data value; as a platform operator, it can quickly seize the first opportunity in the open data market to enhance industrial influence and discover a new development path.

- Meet the need for secure data convergence: By keeping the data within the financial institution, PEC can enable credit risk control and improve precise marketing efficiency with multi-party data.
- Provide precise marketing: PEC can enable more accurate recognition and analysis of customer needs for precise marketing and better customer experience, which can improve overall competitiveness.
- Enhance risk control management capability: PEC can bring more efficient and reliable risk management through full aggravation analysis within data integration from other parties to detect risky fraud groups, which can strengthen risk prediction and control capabilities with fewer risk control staff.

## 9.3 Precise marketing for an automobile enterprise

This chapter introduces the use case of " Telco+Automobile" in the joint marketing scenario.

### 9.3.1 Client requirements

Precise marketing is an essential part of the automobile market competition. Traditional marketing recommendation is constructed based on data owned by automobile enterprises, with limitations such as low timeliness and accuracy, as well as inadequate data dimensions and data sample sizes, resulting in insufficient modeling precision and inefficient follow-up marketing with wasted human resources, as well as missing business opportunities. Driven by multiple factors such as the accelerated digital transformation of automobile enterprises and increased data security requirements, traditional automobile enterprises urgently need secure and reliable innovation marketing paths.

### 9.3.2 Solution and effects

An automobile enterprise seeks external data for cross-domain cooperation and empowerment; while ensuring data privacy security for all parties, AsiaInfo's solution has assisted the enterprise in recognizing customers with high repeat/replacement purchase intent, and bridged links between marketing segments for repeat/replacement purchase.

This solution combines the advantages of industrial data from both the telco and the automobile enterprise with continuous model reasoning and real-time prediction of customers and purchase intentions. The joint analysis under non-equilibrium conditions involves the analysis and modeling of 1.3 billion data and more than 1,000 model labels on the telco side. Based on the reasoning results of actual samples, the model shows good prediction capability to support application touch.

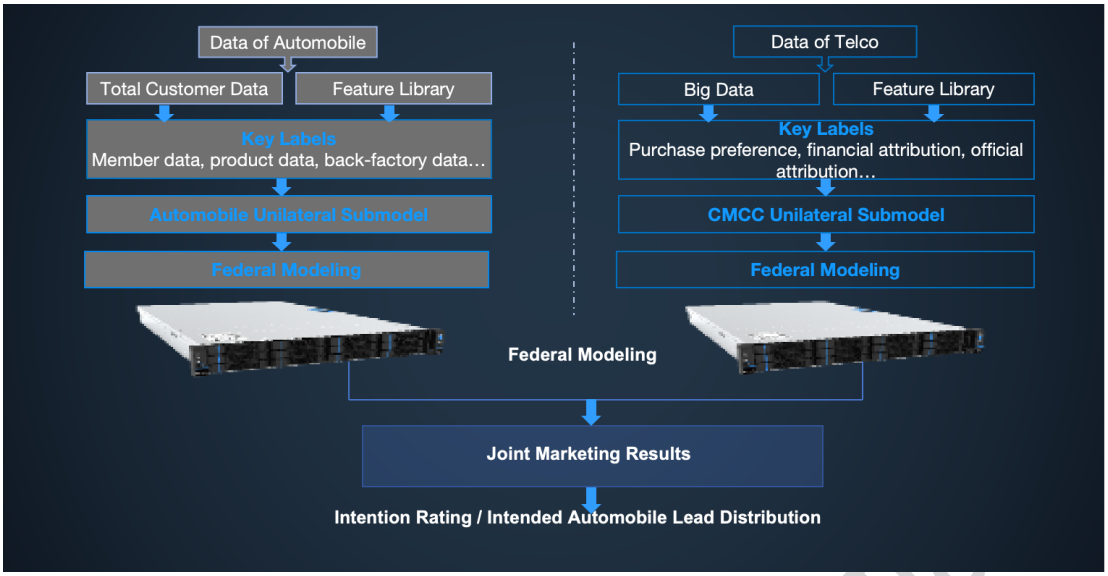



**Figure 9-7 Example of modeling process**

Through the implementation, the marketing activities have been evaluated in a closed-loop with significantly improved arrival rate, coupon rate and effective clue percentage of repeat/exchange marketing, reflected in:

- Increased repeat/exchange purchase intent rate by 60% and AUC by 20%;
- Increased customer coupon rates by 30% and store visits by 56%;
- Increased the percentage of effective referral clues by 68% and over 20% savings in marketing costs.


**Figure 9-8 Use Case effects**

## 9.4 Intelligent Recommendation for medical care

This chapter introduces the use case of "Telco+Medical" in Intelligent Recommendation scenario.

### 9.4.1 Client requirements

As a group-wide benchmark and the largest telecom operator in the province, a provincial operator faces the following pain points in data value mining and model enhancement:

- Large data assets cannot be realized due to privacy regulations.
- Model accuracy enhancement encounters bottlenecks because of insufficient data dimensions and data volume, which is in need to improve model performance.
- It is unable to realize massive training and connectivity, and data volumes such as IoT and 5G are particularly huge without a convenient aggregation method for training.

The client needs to conduct joint modeling of multi-party data and collaborative data utilization to empower intelligent recommendation scenarios based on data privacy-preserving and security compliance between the operator and the partner medical institution. The main requirements include:

- Capable to cross-domain model within data privacy security limits;
- Capable to provide extensible collaborative modeling, and joint learning among multiple parties;
- Equipped with system management functions, including but not limited to project management, organization management, user management, role management, and other related functions.

### 9.4.2 Solution and effects

AsiaInfo has developed three models for its client. The first is an intelligent recommendation model for "Telco+Medical" based on vertical FL through a new distributed machine learning paradigm; the second is an intelligent recommendation model for the medical industry based on the FL framework to achieve localized deployment and co-coordination between the operator and the partner medical institution; and the third is a federation recommendation model for expert consulting and speedy consultation based on the data characteristics

from operator and partner medical institution, which is applied on the client's appointment App.

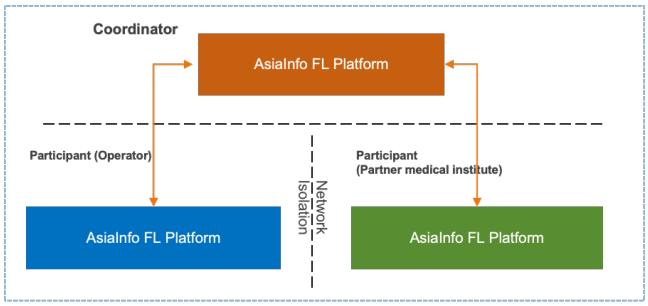



**Figure 9-8 Deployment scenario diagram**

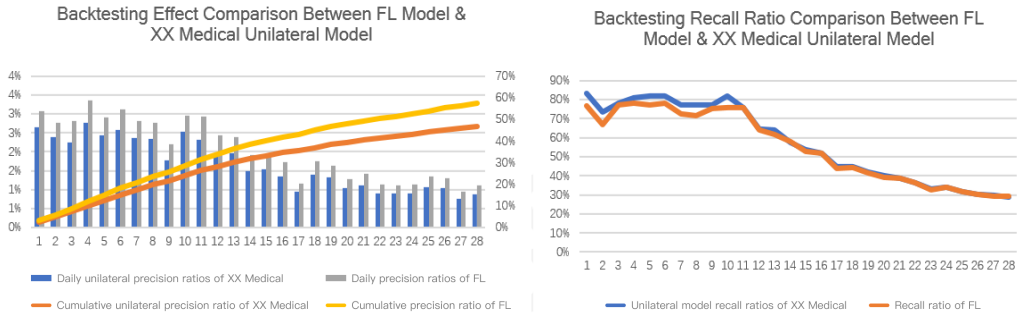



**Figure 9-9 Effects of Intelligent Recommendation with FL model**

With Intelligent Recommendation models based on vertical FL, the following effects have been achieved:

- Establish FL model architecture for the virtual data bridging between the operator and the partner medical institutions under the premise of data not outbound;
- Create a federal recommendation model for expert consulting/speedy consultation for precise marketing and solutions to the same content recommended to all users, with a 10% increase in click rate, a 50% increase in conversion rate, and a 10% increase in accumulated accuracy checking rate;
- Gradually expand the recommendation scenarios such as medical examination, registration, and doctor recommendations, which eventually realize customized recommendations for different users logging into the APP by federal mobile data.

# 10 Certificates and Awards

The use case of AISWare PEC in the automobile industry was selected in State of Privacy-Preserving Technologies in Asia Pacific by Forrester in 2023.


**Figure 10-1 Selected in *State of Privacy-Preserving Technologies in Asia Pacific* by Forrester**

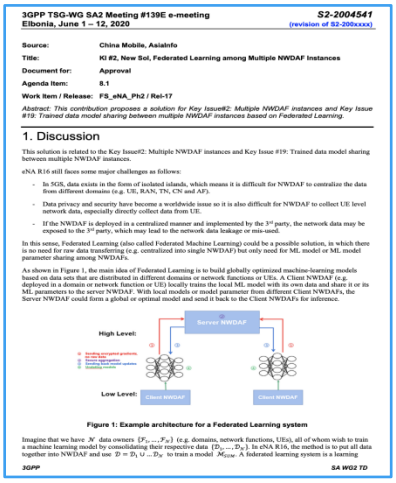
AsiaInfo introduces privacy-preserving computation to the TMF architecture framework for the first time, which has been selected as a TMF Catalyst Outstanding Project.

**Figure 10-2 Contribution to TMF System Standards**

AsiaInfo serves as the leader of the following standards:

- IEEE P3117™ - Draft Standard for Interworking Framework for Privacy-Preserving Computation
- IEEE P3127 Blockchain-based Federated Machine Learning
- IEEE P2986 Recommended Practice for Privacy and Security for Federated Machine Learning

The new FL technology standard for 5G core network element NWDAF proposed by AsiaInfo has been adopted by 3GPP, the world's leading standardization organization.



3GPP TSG-WG SA2 Meeting #139E e-meeting
Elbonia, June 1 – 12, 2020

S2-2004541
(revision of S2-2005xxx)

Source: China Mobile, AsiaInfo
Title: KI #2, New Sol, Federated Learning among Multiple NWDAF Instances
Document for: Approval
Agenda Item: 8.1
Work Item / Release: FS_eNA_Ph2 / Rel-17

*Abstract: This contribution proposes a solution for Key Issue#2: Multiple NWDAF instances and Key Issue #19: Trained data model sharing between multiple NWDAF instances based on Federated Learning.*

1. Discussion

This solution is related to the Key Issue#2: Multiple NWDAF instances and Key Issue #19: Trained data model sharing between multiple NWDAF instances.

eNA R16 still faces some major challenges as follows:

- In 5GS, data exists in the form of isolated islands, which means it is difficult for NWDAF to centralize the data from different domains (e.g. UE, RAN, TN, CN and AF).
- Data privacy and security have become a worldwide issue so it is also difficult for NWDAF to collect UE level network data, especially directly collect data from UE.
- If the NWDAF is deployed in a centralized manner and implemented by the 3rd party, the network data may be exposed to the 3rd party, which may lead to the network data leakage or mis-used.

In this sense, Federated Learning (also called Federated Machine Learning) could be a possible solution, in which there is no need for raw data transferring (e.g. centralized into single NWDAF) but only need for ML model or ML model parameter sharing among NWDAFs.

As shown in Figure 1, the main idea of Federated Learning is to build globally optimized machine-learning models based on data sets that are distributed in different domains or network functions or UEs. A Client NWDAF (e.g. deployed in a domain or network function or UE) locally trains the local ML model with its own data and share it or its ML parameters to the server NWDAF. With local models or model parameter from different Client NWDAFs, the Server NWDAF could form a global or optimal model and send it back to the Client NWDAFs for inference.



Figure 1: Example architecture for a Federated Learning system

Imagine that we have $\mathcal{N}$ data owners $\{\mathcal{F}_1, ..., \mathcal{F}_N\}$ (e.g. domains, network functions, UEs), all of whom wish to train a machine learning model by consolidating their respective data $\{\mathcal{D}_1, ..., \mathcal{D}_N\}$. In eNA R16, the method is to put all data together into NWDAF and use $\mathcal{D} = \mathcal{D}_1 \cup ... \mathcal{D}_N$ to train a model $\mathcal{M}_{SUM}$. A federated learning system is a learning

3GPP SA WG2 TD

**Figure 10-3 3GPP TSG-WG SA2 / Rel-17**

# 11 Contact Us

**AsiaInfo Technologies (China) Limited**

**Address:** AsiaInfo Plaza, Coutyard#10 East, Zhongguancun Software Park Phase II, Xibeiwang East Road, Haidian District, Beijing, P.R.China

**Postcode**: 100193

**Fax:** (+86) 010-82166699

**Tel:** (+86) 010-82166688

**Email:** 5G@asiainfo.com

**Web:** www.asiainfo.com


# Thank you

Customer Value Innovator & Digital Transformation Promoter with Full-Stack Data Intelligence Capabilities