Research on Construction and Key Technology of Water Conservancy Data Middle Platform

Hongsheng Qian1,2*, Liming Xu1,2, Chi Zhou1,2, Zhizhang Qiu1,2
1 Zhejiang Institute Of Hydraulics & Estuary (Zhejiang Institute of Marine Planning and Design), Hangzhou, Zhejiang 310020, China
2 Hangzhou Ding Chuan Information Technology Co., Ltd, Hangzhou, Zhejiang 310020, China
*Corresponding author’s e-mail: qian.hongsheng@163.com

Abstract. The smart water conservancy is an important part of Digital China and Smart Society and an inevitable requirement of the current water conservancy informationization development, the water conservancy data middle platform is a key part of the smart water conservancy. This article builds an overall technical framework as well as studies and researches the key technologies of water conservancy data aggregation, water conservancy data development, water conservancy data governance and water conservancy data service, on the basis of the general technical framework of “four horizontal and three vertical” in Zhejiang provincial water conservancy digital transformation, so as to provide a good theoretical reference for construction of water conservancy industry data middle platform of Zhejiang Province.

1. Preface
In the report of the 19th National Congress of the Communist Party of China, it is clearly proposed to use Internet technology and information technology to build a cyber power, a digital China and a smart society. In December 2018, the Zhejiang Provincial Government issued Zhejiang Provincial Overall Plan on Deepening the ‘Run Once At Most’ Reformation and Promoting the Government's Digitization Transformation[1], which upgraded the modernization of government governance system and governance capacity. In July 2019, the Ministry of Water Resources issued Three-Year Action Plan for the Improvement of Water Conservancy Network and Information Level, the Guidance Opinions on Accelerating the Promotion of Smart Water Conservancy, The Overall Plan for Smart Water Conservancy and other documents [2-3], which clarified the general requirements and main tasks of smart water conservancy in the coming period. Smart water conservancy is an important part of digital China and smart society, and the water conservancy data middle platform is the core functional module of smart water conservancy.

Zhejiang Provincial Department of Water Conservancy has started the digitization transformation of water conservancy from 2019, in accordance with the provincial government's overall technical framework of "Four Horizontal and Three Vertical" for digitization transformation and the overall layout of "One Cloud, One Warehouse and One Platform" for the whole industry, focusing on "business collaboration and data sharing", has established the construction norms of “Five Unification”, thus unified user, unified portal, unified data, unified map and unified security, has created a unified working platform for provincial, municipal and county water conservancy
departments\(^{[4-6]}\), and has achieved certain achievements. However, as it is the most critical in the
digitization transformation of water conservancy, water conservancy data middle platform
construction is still in the stage of exploration at present.

2. Construction Objective

Through constructing the water conservancy date middle platform, carding comprehensively the water
conservancy core business, to achieve business collaboration and data sharing, to realize “on date, one
source, one responsibility” of water conservancy data, to build a unified provincial, municipal and
county platform of "business cooperation, decision support, data sharing, online Internet"\(^{[6]}\), to break
the physical island of water conservancy data, so as to drive the modernization of Zhejiang Provincial
water conservancy management system and management capacity.

The objectives of building water conservancy data middle platform are as following:

- Overall business collaboration. Applications systems development status is transformed from
"chimney mode" into business comprehensive coordination, so as to achieve the overall high
efficiency across applications systems, departments, and levels.

- Scientific decision making. The decision making support model is changed from information
monitoring and convergence display to accurate prediction and intelligent decision making, so as to
comprehensively support strong supervision of water conservancy industry.

- Overall data sharing. Data resources are transformed from dedicated and isolated to
code, sharing and mutual integration, so that thorough perception and cross-application
system, cross-department sharing, cross-level sharing is realized.

- Real-time online service. The application service mode has changed from "preliminary
online and partially on mobile" to "comprehensive online and basically on mobile", realizing the full
coverage of "online" and "on mobile" of the core water conservancy business.

3. Overall Architecture

The overall architecture of water conservancy data middle platform is shown in Figure 1. The water
conservancy intelligent perception layer provides the data basis for the water conservancy data center,
and the water conservancy data center provides water conservancy data service for specific water
conservancy business applications. The water conservancy data center plays a key role in the
architecture.
3.1. Water Conservancy Intelligence Perception

Water Conservancy intelligence perception is a three-dimensional perception monitoring network based on Internet of Things of water conservancy and the integration of heaven and earth. Based on the water cycles as a clue, the three-dimensional and efficient monitoring system of water resources, water environment, water ecology, water security, water management and other information is realized. The perception system mainly includes:

1. **Space based**: Space Satellite Observation Platforms, such as cloud water on Terra/Aqua, precipitation on GPM/TRMM, underground water on GRACE, soil moisture on SMAP;
2. **Ground based**: Ground observation platforms, such as weather station, hydrologic station, water quality station, ecologic station, flow station, underground water station;
3. **Sky based**: Aerial observation platforms, such as cloud water, precipitation, evapotranspiration soil moisture;
4. **Internet based**: Internet platforms, such as wading public opinion, industrial production.

With the continuous development of new technologies, the use of satellites, radar, unmanned aerial vehicles, video, remote-controlled ships, robots, as well as 5G, NB-IoT and other new generation of communication technologies, further coverage of the perception range can be achieved and a broader water conservancy intelligence perception network can be built.

3.2. Water Conservancy data middle platform

Water conservancy data middle platform is the core part of smart water conservancy, which mainly includes the following parts:

1. Water data aggregation. Water conservancy data aggregation is the first link of water conservancy data middle platform construction. Its main purpose is to break the physical island of water conservancy data, to form a unified water conservancy data center, and provide original materials for the subsequent construction of water conservancy data middle platform. Water conservancy data aggregation is to extract, collect, integrate and process water conservancy data,
collect heterogeneous water conservancy data to a unified platform for storage, and then process water conservancy data through modeling, and turn it into useful data for business. Water conservancy data aggregation methods generally include database synchronization, buried point, web crawler, message queue. ETL tools, such as Kettle, Datastage, Datax, Kafka, Oracle Goldengate, are generally used for data aggregation. Through these tools, water conservancy data are classified and aggregated into various basic databases, monitoring databases, business databases, subject databases, spatial databases and other databases, providing a basis for the next step of data development and governance.

(2) Water conservancy data development. Water conservancy data development is to transform data into data assets, into new forms of business needs, and extract data value. 4V characteristic of big data determines the data processing is a complicated project. The construction of water conservancy data need to build the infrastructure tools of the data middle platform, to meet all kinds of structured and unstructured data collection, storage and processing, to deal with off-line and real-time computing and storage of data in line with the scene, and ensure the operation of the data can be assigned to the business side after each data processing task is chained together with another. Water conservancy data development includes offline development, online development and data model. Offline development includes offline tasks, resource scheduling, task monitoring, task operation and maintenance. Offline development computing tools mainly include Spark, Hive, scheduling platforms Azkaban, Oozie, EasySchedule. Online development includes real-time task, resource scheduling, task monitoring, task operation and maintenance. Online development tools mainly include Flink, SparkStructStreaming, SparkStreaming and Storm. Data models include algorithm development, model training, model evaluation, model monitoring. Common development tools include TensorFlow, Spark ML.

(3) Water conservancy data governance. Data governance is a process to evaluate the quality of the integrated massive multi-source heterogeneous data and govern the original scattered, repeated and low-quality data into a high-quality data set with unified format, type, unit, consistent coding and logic, and clear number source through the customized development of data extraction, cleaning, conversion, fusion and loading processes, which includes three parts: data assets, data quality and data security. Data assets mainly establish corresponding data standards and specifications by comprehensively combing the content of core data assets with the help of technical means and tools. Data assets mainly establish corresponding data standards and specifications by comprehensively combing the content of core data assets, with the help of technical means and tools, effectively implement the definition standards and specifications, including metadata management, data graph, data consanguinity, data categories. Data quality and data security run through the whole data governance stage: Data quality is usually evaluated from six dimensions, including completeness, standardization, consistency, accuracy, uniqueness, and timeliness, which is mainly realized through data cleaning, ID Mapping and other technologies, including data verification, multi-source verification, data reporting. Data security is to ensure the security of data assets in the using and sharing exchange process, combined with the technical means of information security, including data desensitization, data encryption, rights control, and to realize the desensitization, encryption and decryption of data, rights management, according to the organization, role, post, users, etc.

(4) Water conservancy data services. Water conservancy data service is a service capacity, which is transformed from the data assets, to encapsulate the high versatility of business and algorithm into a micro service, according to the technology architecture of "small foreground, big middle platform" proposed by Alibaba in 2015, to establish water conservancy business middle platform. The water conservancy business middle platform mainly including general basic component micro-service such as information retrieval, message send, task scheduling, process engine, rule engine, real-time data interface micro-service for monitoring data such as water level, rainfall, flow, water quality, and water intake and business smart micro-service including flood forecast and early warning, typhoon route prediction, basin flood joint dispatching, regional water resources carrying capacity early warning, and abnormal prediction of dynamic monitoring in water area. The water conservancy data service is mainly built on the basis of Spring Boot, Spring Cloud, Spring Security, etc. It uses intelligence
perception, big data mining, artificial intelligence AI recognition, cloud computing, block chain and other advanced intelligent algorithms, and uses a unified service gateway to provide water conservancy data service for various water conservancy business applications.

3.3. Water Conservancy Business Application
The water conservancy business application layer mainly realizes the application of water conservancy intelligence. Centering on the seven aspects, thus, water disaster prevention, river, lake and reservoir protection, water resources security, water development planning, water affairs supervision, water administration coordination and public service, following the steps of comprehensive perception, intelligent analysis, prediction and early warning, decision making support, and on the basis of existing informatization, the application of business, affairs and government affairs is deepened. Through BI report, visual large screen, data mining platform and other tools, the water conservancy project construction and operation management and collaborative office and other application systems are carried out, so that it is achieved the water conservancy performance behavior process can be traced and responsibility can be investigated, and the water conservancy industry digital and intelligent level is significantly improved.

4. Conclusion
The smart water conservancy is an important part of Digital China and Smart Society and an inevitable requirement of the current water conservancy informationization development, and the water conservancy data middle platform is the key component of smart water conservancy. Complying with technical framework of "Four Horizons and Three Verticals" for Zhejiang Provincial digitization transformation of water conservancy and the overall layout of the whole industry "One Cloud, One Warehouse and One Platform", this paper constructs the water conservancy data middle platform, studies and explores the key technologies such as water conservancy data aggregation, water conservancy data development, water conservancy data governance and water conservancy data service.

At present, Zhejiang Province has carried out some work in water conservancy data aggregation and water conservancy data service. Some achievements have also been made. By the end of January 2021, it had collected 320 million pieces of water conservancy data, developed a data supermarket platform and a data sharing and exchange platform, and provided more than 1,000 water conservancy data services, which had been called 130 million times. In the next step, more theoretical research and development will be carried out in terms of water conservancy data development and water conservancy data governance to continuously enrich and develop the water conservancy data middle platform in practice, so as to provide strong support for the construction of Zhejiang Provincial smart water conservancy.

References
[1] Zhejiang Provincial People's Government. (2018) Notice of Zhejiang Provincial People's Government on Printing and Issuing the General Plan of Zhejiang Province to Deepen the Reform of "Run at Most Once" and Promote the Government's Digital Transformation. Zhejiang Provincial People's Government, 1:2-13.
[2] Ministry of Water Resources of the People's Republic of China. (2019) Notice of the Ministry of Water Resources on Printing and Issuing the Three-Year Action Plan for Improving the Water Resources Network Informationization Level (2019-2021). Ministry of Water Resources, PRC, 1:1-32.
[3] Ministry of Water Resources of the People's Republic of China. (2019) Notice of the Ministry of Water Resources on Printing and Issuing the Guidance Opinions on Accelerating the Promotion of Smart Water Resources and the Overall Plan of Smart Water Resources. Ministry of Water Resources, PRC, 1:1-100.
[4] Office of Water Conservancy Department of Zhejiang Province. (2019) Notice of the Office of Water Conservancy Department of Zhejiang Province on Printing and issuing of the General Plan of Water Management Platform in Zhejiang Province. Office of Water Resources Department of Zhejiang Province, 1:1-92.

[5] Yu KS, Jiang XJ, Jin XC. (2018) Smart Water Conservancy Plan of Zhejiang Province under the Digital Transformation of Government. Water Resources Informatization, 6: 11-15.

[6] Bao ZY, Jiang XJ, Huang K, et. (2020) Research on the General Framework and Key Technologies of Digital Transformation of Zhejiang Water Resources. Water Resources Informatization, 2: 1-8.