

Multi Dimensional and Scenario Generalization Value Quantification of Data Assets Based on a Multi Factor Valuation Scoring Model — Taking the Water Resources Data Assets of Key Provinces in the Comprehensive Management of the Yangtze River Basin as an Example

Fu Rao^{1,a,*}

¹Department of Finance, Zhejiang University of Finance & Economics, Hangzhou, China
a. 3347695482@qq.com

*corresponding author

Abstract: Nowadays the explosive growth of digital assets has become a new economic growth point. The entry of data assets into the balance sheet and embracing the capital market to exert its multiplier effect have become trending topics. However, the quantification of data asset valuation, as data asset utilization's cornerstone, still faces many obstacles, such as discriminatory analysis, lack of scenario generalization valuation systems and multi objective quantitative dimension, etc. Starting from both subjective and objective dimensions and coordinating multiple market entities, this article constructs a multi factor valuation screening and scoring model to quantify the value of data assets and addresses these practical challenges. Quantify factors through methods such as the Short Ratio Method and Delphi Expert Method. Due to the heterogeneous preferences and requirements of different industries for different factors, this article uses the Compound Analytic Hierarchy Process and effect functions to rank and combine factors for weighting, meeting the specific application effects in various scenarios while also achieving a scenario generalization approach to meet the needs of cross industry trading of data assets. Finally, apply the model to the water resources data assets of the Yangtze River comprehensive management province. This case study promotes the valuation of water resources data assets and the multi-party trading and circulation between governments and market entities. This article aims to provide new models and academic recommendations for data asset valuation and value quantification, to assist in the utilization and circulation of data assets, as well as the construction of data asset trading markets.

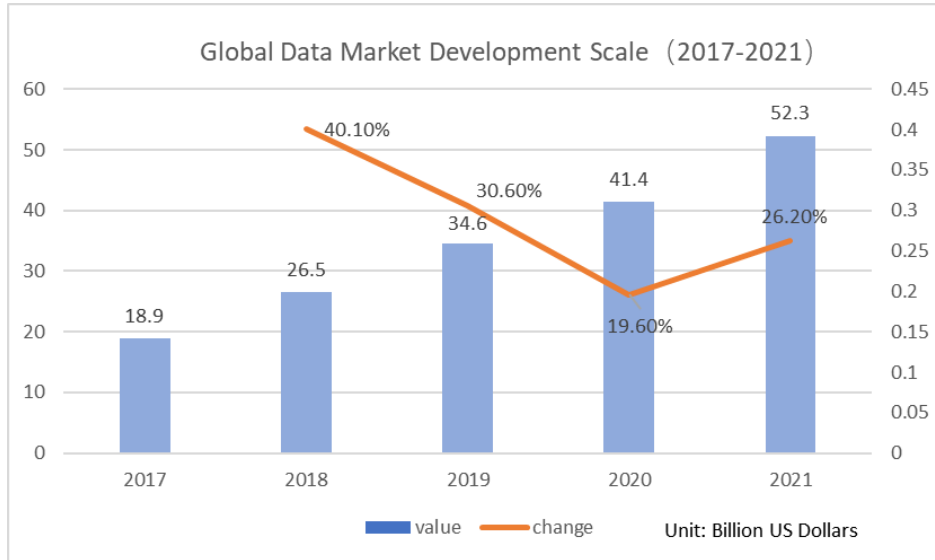
Keywords: quantitative valuation of data assets, multi factor valuation screening scoring model, Delphi Expert Method, Compound Analytic Hierarchy Process, factor sorting and weighting.

1. Introduction

The "Digital China Development Report (2023)" released by the National Data Administration on May 24, 2024 shows that the added value of core industries in the digital economy will account for 10% of GDP in 2023, and the data factor market will become increasingly active. The total data

production in 2023 will reach 32.85ZB, a year-on-year increase of 22.44%; According to IDC data, the compound annual growth rate of the big data market is expected to reach 10.4% globally over the five-year period from 2020 to 2024.

Data, as the fifth major production factor, has rapidly integrated into various links such as production, circulation, and consumption, evolving from a byproduct of economic cycles to a new engine driving high-quality economic and social development. With “the fourth wave of artificial intelligence”, the scale of data assets has grown exponentially, and the evolution of assetization is significant. Its economic value and development potential are also increasingly prominent.



(Source: Research on Industrial Transformation)

Figure 1: PXR Global Data Development Scale Statistics from 2017 to 2021.

To further promote the appreciation and revenue generation of data assets and amplify the multiplier effect of data assets, promoting the financialization of data assets is imperative. In December 2023, the National Data Administration and 17 other departments jointly issued the "Three Year Action Plan for 'Data Elements ×' (2024-2026)" to strengthen the traction of scenario demand, drive high-quality supply and efficient flow of data elements in compliance.

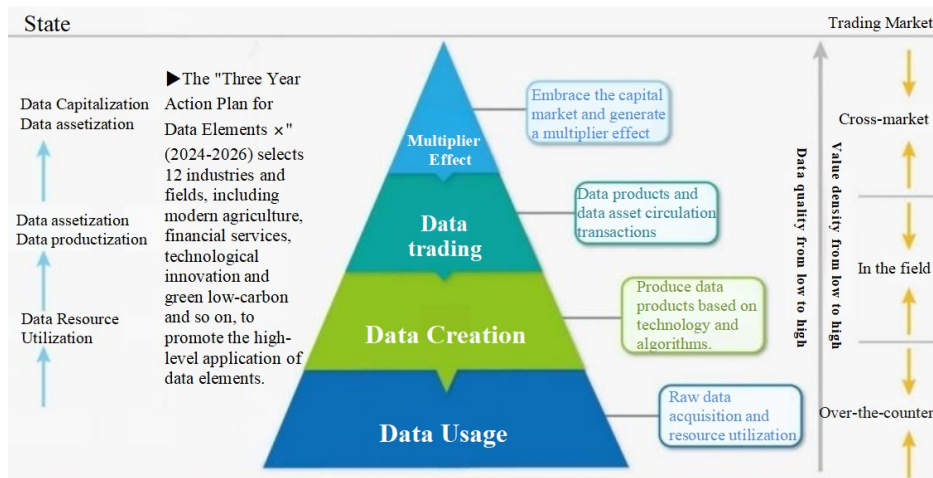


Figure 2: The Pyramid Shaped Marketization Process of "Data Usage → Data Creation → Data trading → Multiplier Effect" for data assets.

However, the stimulation of the economic value of data assets cannot be separated from data governance work such as accurate valuation, value quantification, and ownership and pricing of data assets. Currently, data assets are in their infancy, and there is an urgent need for comprehensive, multi-dimensional, and scenario based valuation and value quantification methods with excellent application effects for data assets.

2. Literature review

2.1. Realistic Challenges and Research Status

2.1.1. Difficulty in Valuing Data Assets

Currently, the development and utilization of data assets is still in its early stages globally, and there are three major challenges in terms of valuation.

Firstly, the subjective efficacy rate is low. In the process of data quality assessment, manual evaluation is the main method, which has high evaluation costs, low efficiency, and significant subjectivity. This leads to inconsistent control scales in data quality assessment for data exchanges, hindering the free flow of data elements [1][2].

Early scholars mainly focused on subjective valuation and minority quantitative valuation in the pricing and valuation of data assets, and there was also a significant subjectivity in determining indicator parameters. For example, Even and Shankaranarayanan use utility driven theory to measure data quality by the ratio of actual utility to potential maximum effect [3]. However, the determination of its utility value requires manual evaluation, which is subjective and inefficient. The "Data Potential Energy" valuation system constructed by PwC in the "White Paper on Open Data Asset Valuation" in 2021 also has a strong subjectivity in quantifying indicators based entirely on manual scoring.

With the development of computer technology, many scholars began to approach objective quantification in later explorations. For example, by transmitting the perturbed image to the model to calculate perturbation entropy, perturbation entropy can distinguish good data from bad data from the perspective of information value, and consider it as the image data quality score [4]. And some scholars evaluate data quality from the perspective of information content and use distance entropy method to screen data quality [5]. However, there are still certain issues with the interpretability of these methods for data quality, as they only quantitatively evaluate data quality from a single or relatively limited perspective.

Although current research has improved the application efficiency of data asset valuation and pricing, existing data quality assessment methods face the dilemma of being inadequate in the face of explosive data growth and more complex market environments. It is urgent to find a value quantification evaluation model that integrates subjective and objective multidimensional factors, coordinates buyers and sellers, as well as third-party institutions and other market entities.

2.1.2. Difficulty Balancing Meeting Application Scenario Requirements with Generalized Valuation System

In the early days, data quality assessment methods mostly focused on comprehensive data governance and often overlooked data application scenarios. This poses a challenge to determining the transaction price of data elements, resulting in lower willingness of buyers to pay in data transactions [6]. The commonly used evaluation methods include the comprehensive evaluation method, which mainly includes the fuzzy comprehensive evaluation method fuzzy hierarchy [7], analytic hierarchy process [8], Delphi method, etc. Among them, although these universal data quality assessment methods have strong applicability, their specific application effects are still unsatisfactory [9].

To alleviate this difficulty, scholars have begun to study characteristic valuation methods for specific application scenarios to improve their practical usability [10]. For example, regarding government data assets, scholar Song Mufan proposed in 2022 a government open data quality evaluation index system based on four dimensions: data content, data expression, data acquisition, and data utility [11].

Although personalized evaluation systems have to some extent improved the practicality of data asset valuation models, current data often flows between industries, and various industries are forming interconnected information ecosystems. Specialized valuation models often fail to find a consensus on balancing the valuation of data assets across industries, thus unable to meet the needs of cross industry data asset transactions and circulation.

In addition, with the rapid development of the current industry and the emergence of many emerging industries, setting a valuation system for each industry one by one will obviously lead to the dilemma of "supply shortage", which cannot adapt to the rapid development and accelerated changes of the times.

Therefore, it is urgent to propose a data quality assessment method that can meet both application scenarios and wide applicability, in order to meet the new needs of data growth.

2.1.3. Discriminatory Analysis and Data Asset Foam

Li Yaxiong et al. pointed out that data assets are measurable resources that can be processed by enterprises to achieve specific business purposes and bring economic benefits. This means that the raw data itself does not constitute the assets of the enterprise, only the data processed by algorithms has real value [12]. However, in order to achieve specific interest demands and profit maximization goals, enterprises often implant their own value orientation and interest intentions when developing algorithms [13], which will result in discriminatory analysis of data assets.

Discriminatory analysis of data assets will lead buyers of data transactions into unfair trade traps, hinder the equal and free flow of data elements, and be detrimental to the long-term stable and sustainable development of data asset transactions. At the same time, discriminatory analysis can make it difficult for society to accurately value a company's data assets, making it difficult for all parties in society to fully utilize the company's data assets. Low value and low-quality data can also cause practical problems such as reduced efficiency and waste of social resources. Because data prediction results often rely on the data itself, improving data quality can effectively reduce prediction errors [14]. At the same time, discriminatory analysis will also cause the butterfly effect of "data asset foam".

2.1.4. The quantification of digital asset value is not yet unified, and the commercialization transaction rate and return rate are low

Due to difficulties in valuing data assets, identifying ownership, and conducting discriminatory analysis, there is currently no social consensus on the valuation of data assets. The evaluation and measurement of the value of numerous data assets are still in the qualitative stage, with a lack of quantification, and the data asset circulation and trading market has not yet taken shape. Zhao Zhigang, director of the China Academy of Fiscal Sciences, said, "Data asset financing still faces many practical challenges. The unstable value of data assets, the lack of a mature and active data trading market, and the easy loss of data asset value make it difficult to pledge a large amount of data at present.

In 2023, the Shanghai Data Exchange released the "Shanghai Roadmap for Data Asset Securitization", proposing a "one bridge, two institutes, two axes" architecture for data asset infrastructure DCB (Data Capital Bridge). This coincides with one of the core objectives of this article,

which is to transform data assets into tradable tokens, so that data assets can truly circulate in the trading market like stocks and bonds in the future, achieving asset realization.

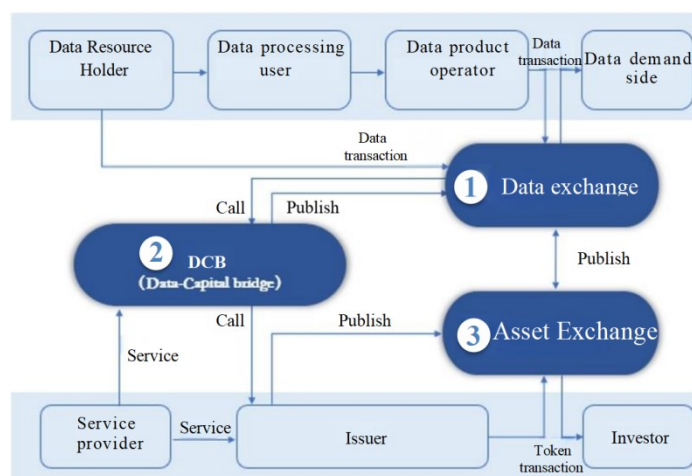


Figure 3: DCB data asset "one bridge, two institutes, two axes" architecture.

However, the process of commercializing data assets cannot be achieved without building a unified and standardized value evaluation system for data assets, which is also a prerequisite for improving the trading rate and return rate of data assets.

2.2. Practical Significance

Firstly, this study is of great significance for enterprises, especially technology-based and small and micro enterprises that prioritize 'light assets and heavy data'. In the first quarter of 2024, 18 listed companies disclosed data assets in their financial reports, involving a total amount of 103 million yuan. The impact of 'entering the table' on the profit margins of two companies exceeded 10%; Bringing a theoretical market value increase of over 10 million yuan to a listed company, exceeding 1% of the total market value of the enterprise.

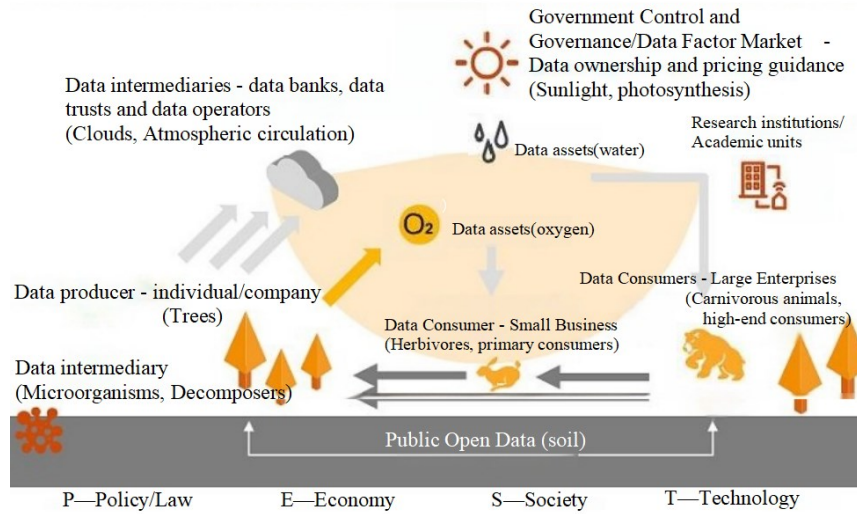
From this, it can be seen that data assets will become a new value growth point for future enterprises in terms of revenue generation, financing, and appreciation, with enormous revenue generating potential. As Professor Li Feng from Shanghai Jiao Tong University said, "Data asset financing can activate existing assets, unleash the economic value of data, not only help these enterprises solve financing problems, but also promote the development of the digital economy and accelerate the formation of new quality productivity."

Therefore, the research and exploration of multi-dimensional, multi-party, and pan scenario value quantification models for data assets can help enterprises accurately value and improve the utilization and return of data assets.

Secondly, exploring the essence of data asset value evaluation and constructing such a value quantification model is conducive to the integration and sharing of data assets across multiple fields in society, reducing discriminatory valuation and maximizing social value.

3. Construction of a Multi Factor Valuation Screening Scoring Model

The market where data assets can be fully utilized and coordinated for circulation should be a collaborative, multi factor, harmonious, standardized, and secure data asset ecosystem. This also puts forward multi-dimensional, multi factor, and multi-level requirements for the construction of our valuation model.



(Source: PwC Analysis)

Figure 4: Data Asset Ecosystem.

To construct a data asset value quantification system that integrates objective valuation and subjective scoring dimensions, integrates multiple factors such as enterprises and third-party audits, and is widely used in multiple application scenarios, this paper constructs a multi factor valuation screening scoring model and constructs a three-level factor pool.

3.1. Introduction to Multi Factor Valuation Screening Scoring Model

The multi factor model is a cutting-edge model widely used and performing well in the field of financial quantitative investment [15]. Quantitative analysts evaluate and score stocks within the market by mining effective factors, selecting stocks with good quality and trading performance, and thus achieving valuation and stock selection.

Based on the characteristics of both data assets and stocks, such as large volume, multidimensional factors, and the need for valuation screening, a multi factor valuation screening scoring model is now introduced into the valuation pricing of data assets.

The multi factor valuation scoring model is an effective tool for quantifying the value of data assets. Firstly, multidimensional, multi-party, and multi-level factors are constructed. Then, each factor is sorted and combined with weights based on its preferences. Finally, the quantified values or scoring situations of each factor are weighted and summed to obtain the final quantified value of the data asset value.

Assuming there are n factors, each with a rank of k , the steps for constructing a scoring method for a multi factor model are as follows:

1. List the indicators and reasons for comparing each factor.
2. Explain the ranking method and quantification criteria for each factor.
3. Obtain the ranking, weight, and quantification values of each factor.
4. Calculate the corresponding score for each factor according to the formula, and calculate the total score by weighted summation.

The processing flow of this pair of factors can be applied in the first, second, and third level factor pools. This not only provides a multidimensional comprehensive value assessment and quantification of data assets, but also meets the impact of different application scenarios on the weights of various indicators and factors.

For example, in the financial investment industry, there is a high requirement for the timeliness of data, and the weight of relevant factors can be increased during valuation; For example, in big model training or prediction based on big data, in order to ensure that machine learning produces more accurate and high-quality results, high requirements are placed on data integrity, and the ranking and weight of corresponding factors can be improved.

Thus, at the same time, multi-dimensional data asset valuation scoring and generalized cross industry data asset integration trading have been achieved, promoting the construction of a data asset trading market with accurate valuation, coordinated circulation, cross bank circulation, and value-added income generation.

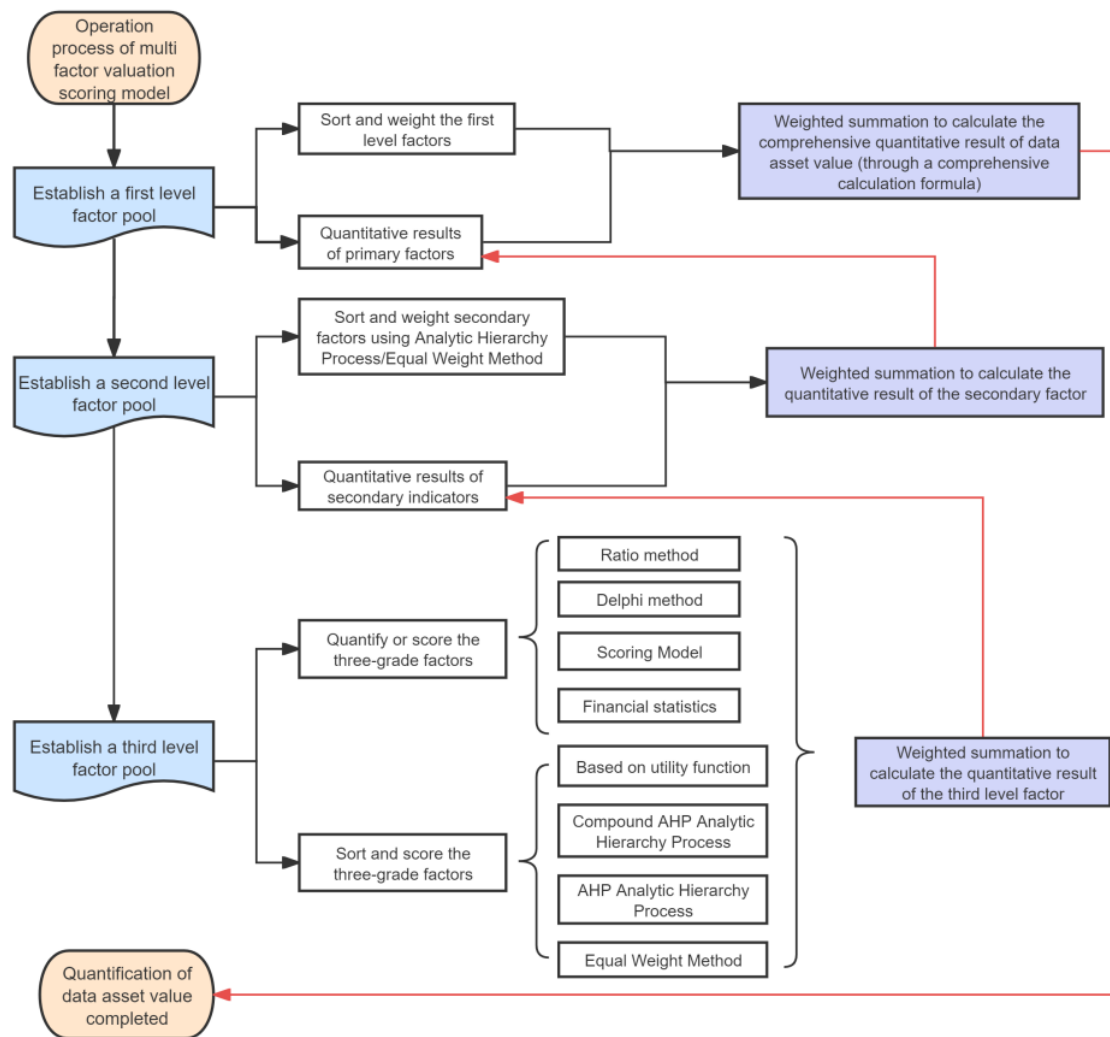


Figure 5: Multi factor Valuation Scoring Model Workflow Diagram.

3.2. Multi factor Model Construction

3.2.1. First Level Factor Pool

This article is based on the interests of multiple parties such as buyers, sellers, and third-party institutions, taking into account both subjective and objective dimensions. Combining the value changes of data assets in the past cost, present value, and future potential, a first level factor pool is established from the perspectives of quality, safety, potential, and cost. These four primary factors

are: objective valuation factor based on utility (V), security factor combined with audit (S), market expectation factor based on AHP analysis (E), and cost value factor (C).

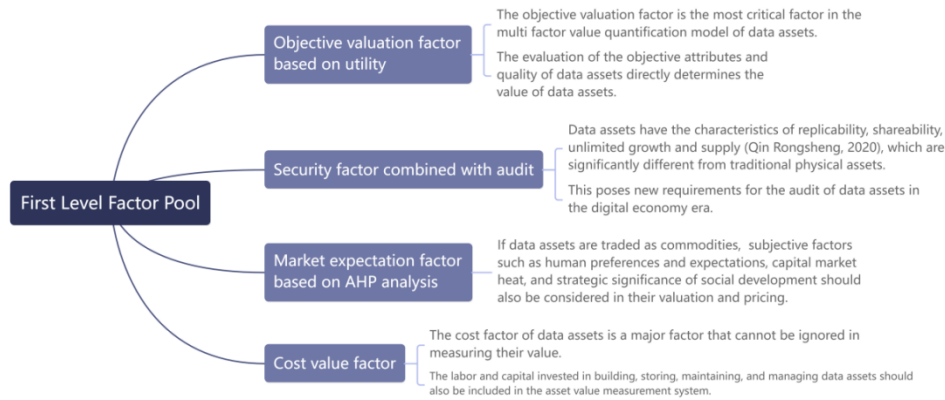


Figure 6: First Level Factor Pool and Reasons for Selecting First Level Factors.

3.2.2. Utility-based Objective Valuation Factor's Second and Third Level Factor Pools

The construction process of this factor pool is to establish second and third level factor pools, analyze the performance of each third level factor in the data, and quantify it. Subsequently, based on the evaluation results of the utility function, each tertiary factor is sorted and weighted. Calculate the quantified total value of the utility based objective valuation primary factor by combining the quantified value and weight summation.

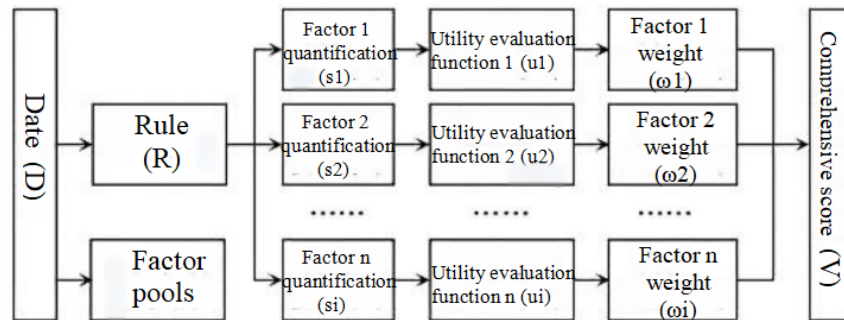


Figure 7: Construction Process of Objective Valuation Factor Pool.

(1)Factor pool construction

This article combines the current "Information Technology Data Quality Evaluation Indicators" (GB/T 36344-2018) standard, "E-commerce Data Asset Evaluation Indicator System" GB/T 37550-2019, and combines with the "White Paper on Open Data Asset Valuation" (PwC, 2021) and the scientific research achievements of scholars such as Tang Yong to determine a system with 13 secondary factors and 25 tertiary factors as the specific evaluation basis. The definitions of each indicator have been elaborated in detail in the relevant documents, and will not be repeated here.

Table 1: Utility-based Objective Valuation Factor's Second and Third Level Factor Pools.

First level factor	Second level factor	Third level factor
	Data Format	Data Carrier
		Data Planning

Table 1: (continued).

Utility-based objective valuation factor		Data Representation
		Data Description
	Accuracy	Language Accuracy
		Abnormal Data Rate
	Integrity	Record Integrity
		Attribute Integrity
	Consistency	Assignment Consistency
		Relationship Consistency
	Timeliness	Update Frequency
		Timeliness of Updates
	Normalization	Standardization of Value Range
		Format Standardization
	Efficiency	Data Duplication Rate
		Use Repetition Rate
	Credibility	Source Credibility
		Method Credibility
	Precision	Data Precision
	Recoverability	Recovery Possibility
		Difficulty Level of Recovery
	Accessibility	Accessibility
	Authenticity	Subjective and Objective Degree
		Proportion of Measured Data
	Data Performance	Data Association
		Data Feature

(2)Factor quantification

The third level factors will be quantified below. s_{ij} is the quantitative value of the third level factor, which is measured using the simple ratio method (Pipino et al.,2002) in this paper. This method is computationally simple and has high robustness. The specific formula is as follows:

$$s_{ij} = \frac{N_{ij}}{M_{ij}} \quad (1)$$

Among them, N_{ij} represents the amount of data in the dataset that meets the requirements of data quality rules, and M_{ij} is the amount of detected data or the value expected by the user. The larger the s_{ij} , the fewer data quality issues and the more characteristic it has.

The quantized value of a secondary factor is equal to the sum of the quantized values of its subordinate tertiary factors.

(3)Factor sorting and weighting based on utility function

Subsequently, we ranked and weighted the secondary factors in the objective valuation factor pool based on utility. Firstly, the utility function is selected to quantitatively evaluate the secondary factors of the detection data.

The utility function can reflect the different requirements of users or all parties for data quality in different application scenarios through risk preference factors. Therefore, the utility function of heterogeneous preferences can quantitatively reflect the tolerance of all parties to different dimensions of data quality issues, so that the final result of quantifying the value of data assets

changes with various real-life application scenarios, that is, to achieve accurate estimation of different value evaluation results for the same data asset in different application scenarios.

The common forms of fitting utility functions include linear function, power function, logarithmic function, and exponential function. Considering that exponential functions can better fit various shapes and characterize preference requirements, the exponential function utility function is selected as the data quality evaluation function. Specifically, as shown below:

$$u_i(s_{ij}) = \frac{1-e^{(a-s_{ij})/\lambda_i}}{1-e^{(a-b)/\lambda_i}} \quad (2)$$

Among them, $u_i(s_{ij})$ refers to the quality evaluation score corresponding to the j -th secondary indicator in the i -th dimension, and the larger its value, the higher the data quality. λ_i reflects the risk preference of data users for the data: when λ_i is less than 0, it is a risk averse type; When λ_i is greater than 0, it is a risk preference type; When λ_i approaches infinity, it is a risk neutral type, which is equivalent to $u_i(s_{ij})=s_{ij}$.

Finally, using the $u_i(s_{ij})$ value as a reference, the secondary factors are weighted and summed to calculate the quantitative value of the objective valuation of the primary factors based on utility. The comprehensive calculation formula is:

$$V = \sum_{i=1}^n s_i * \omega_i \quad (3)$$

V is the comprehensive evaluation value of objective valuation primary factor, n is the total number of effective factors, s_i is the quantified value of each secondary factor, and ω_i is the weight of each secondary factor.

3.2.3. The Second and Third Level Factor Pools of Security Assurance Factor Combined with Auditing

Data assets have the characteristics of replicability, shareability, unlimited growth, and supply, which are significantly different from traditional physical assets. This poses new requirements for the audit of data assets in the digital economy era.

(1) Factor pool construction

The "Requirements for Security Capability of Information Security Technology Big Data Services" (GB/T 35274-2023) states that a data asset operation audit mechanism should be established to achieve auditability and traceability of data asset management operations. Based on the guidance documents on data auditing, such as the "Management Measures for Compliance Audit of Personal Information Protection (Draft for Comments)" and the "Guidelines for Data Compliance Audit", as well as the research results of scholars such as Wang Jingyong on data asset auditing, this article establishes a second and third level factor pool that combines audit security factors, taking into account the principles and practical needs of auditing.

Table 2: The Second and Third Level Factor Pools of Security Assurance Factor Combined with Auditing.

First level factor	Second level factor	Third level factor
	Safety(B1)	Data Access Control (C1)
		Data Leakage Protection (C2)
		Data Environment Security (C3)
	Compliance (B2)	Data Management and Source Compliance (C4)
		Data Usage and Sharing Compliance (C5)

Table 2: (continued).

Security assurance factor combined with auditing	Authenticity (B3)	The actual existence and integrity of data assets (C6)
		Consistency between data asset recording and actual accounting (C7)
		Legitimacy of control over data assets (C8)
		Accuracy of Data Asset Cost Measurement (C9)
		Whether the reporting and disclosure of data assets comply with accounting standards (C10)
	Soundness (B4)	Data Standards and Governance Strategies (C11)
		Data Governance Structure and Organization (C12)
		Data Governance Processes and Practices (C13)

The above factor pool mainly evaluates the security, compliance, authenticity, and soundness of data assets from an audit perspective, starting from four aspects: security, behavior, information, and system.

(2)Factor quantification

Firstly, based on the indicators of each factor in the second and third level factor pools, remove the data that does not meet the standards. However, due to the uneven quality of audit data, we once again use the second and third level factors as indicators to process the compliant data using the Delphi method, in order to more accurately quantify the quality of audit security assurance and lay the foundation for more precise evaluation of the quality and value of data assets in the future.

The first step is to search for multiple audit analysts (6-20 experts) to score the indicators of data assets for the 13 indicators (T1~T13, corresponding to C1~C13 in turn), using a scale of 1-9 points. 1 point indicates the worst degree, and the higher the score, the better;

The second step is to conduct Delphi analysis on the collected and organized data, including conducting Kendall W test on expert scoring data, analyzing the consistency of expert scoring, and visualizing it with a box plot; Calculate the coefficient of variation (CV) value, standard deviation, etc. for concentration analysis. From this, the quantitative scoring results of each three-level factor are obtained.

Table 3: Kendall Coordination Coefficient Range Comparison Table.

Kendall coordination coefficient range: [0,1]				
<0.2	0.2~0.4	0.4~0.6	0.6~0.8	0.8~1.0
The consistency level is poor	The consistency level is moderately low	The consistency level is moderate	The consistency level is strong	The consistency level is extremely strong

Generally, Kendall's coordination coefficient needs to pass the chi square test for analysis. If the $p < 0.05$, it indicates passing and significant consistency. Otherwise, it indicates no significant consistency.

(3) Sorting and weighting of factors

This article uses the Compound AHP Analytic Hierarchy Process and combines it with SPSS AU software to solve and verify the consistency of the weight of this factor.

The hierarchical structure reflects the relationships between factors, and different factors in the target layer (B_i) and indicator layer (C_i) have heterogeneous preferences in the minds of decision makers. Based on the requirements of application scenarios, we scientifically weight them.

Step one, use SPSSAU to construct matrix A and assign weights to $B = \{B_1, \dots, B_n\}$. Compare the impact of n factors on a certain factor A , taking two factors B_i and B_j each time, using B_{ij} to represent the ratio of their impact on A , and using Saaty et al.'s numbers 1-9 and reciprocal as their scales. All comparison results form a comparison matrix A .

$$A = (B_{ij})_{n \times n} = \begin{bmatrix} B_{11} & B_{12} & \cdots & B_{1n} \\ B_{21} & B_{22} & \cdots & B_{2n} \\ \cdots & \cdots & \cdots & \cdots \\ B_{n1} & B_{n2} & \cdots & B_{nn} \end{bmatrix} \quad (4)$$

$b_{ij} \times b_{ji} = 1$.

Step two, use the Analytic Hierarchy Process (AHP) with root mean square method to calculate the weight ω_i of each factor, and select CI and RI as two indicators for consistency testing. If it fails, the judgment matrix will be reconstructed.

Consistency Indicator CI:

$$CI = \frac{\lambda_{\max} - n}{n - 1} \quad (5)$$

The average random consistency index, RI, can be obtained by looking up the table;

Table 4: Random Consistency RI Corresponding Values for Each Order.

n order	3	4	5	6	7	8	9	10	11	12
RI	0.52	0.89	1.12	1.26	1.36	1.41	1.46	1.49	1.52	1.54
n order	13	14	15	16	17	18	19	20	21	22
RI	1.56	1.58	1.59	1.5943	1.6064	1.6133	1.6207	1.6292	1.6358	1.6403
n order	23	24	25	26	27	28	29	30		
RI	1.6462	1.6497	1.6556	1.6587	1.6631	1.6670	1.6693	1.6724		

The consistency ratio CR: $CR = CI/RI$. When $CR < 0.10$, the consistency of the judgment matrix is considered acceptable. Otherwise, appropriate modifications should be made to the judgment matrix.

Step three, using the root mean square method of Analytic Hierarchy Process (AHP) again, construct n judgment matrices B_i , assign weights to the indicator layers C_i under each target layer (B_i), and verify consistency.

Step four, combine the weighting results of the two AHP methods, with a total weight of 100%. By combining the preference ranking of the two layers of factors, more accurate weighting of factors can be achieved to meet the data preferences and requirements of different application scenarios, achieving scenario generalization.

3.2.4. The Second and Third Level Factor Pools of Market Expectation Factor Based on AHP Analysis Method

Data assets with great development potential, high market heat, and key strategic areas can naturally receive higher valuation and pricing. Therefore, this article establishes a market expectation factor for development potential rating and constructs its factor pool. This maximizes the interests of both buyers and sellers, as well as the social benefits of data assets.

(1) Factor pool construction

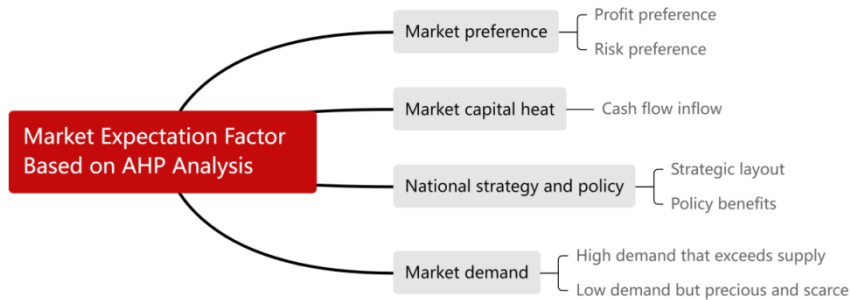


Figure 8: The second and Third Level Factor Pools of Market Expectation Factor Based on AHP Analysis Method.

(2) Factor quantification

The various levels of factors in the market expectation factor pool inherently have subjective attributes, and under the efficient market theory, market sentiment can also have an impact on commodity and asset prices. We use a scoring model to quantify the factors, so that the multi factor valuation screening scoring model takes into account the subjective dimension of expected factors. The following are the specific rules for scoring each factor.

Table 5: Scoring Rules for The Scoring Model.

Factor name	0~20 points	21~40 points	41~60 points	61~80 points	81~100 points
Profit preference	Guaranteed income, with a certain public welfare nature	Yield below 20%	Yield ranges from 20% to 40%	Yield ranges from 40% to 60%	Yield above 60%
Risk preference	Risk averse type	Low risk preference type	Medium risk preference type	Medium to high risk preference type	High risk preference type

Table 5: (continued).

Cash flow inflow	The inflow of main market funds is negative, less than 0	The inflow of main market funds is less than 500 million ¥	The inflow of main market funds ranges from 500 million to 1 billion ¥	The inflow of main market funds ranges from 1 million to 1.5 billion ¥	The inflow of main market funds exceeds 1.5 billion ¥
Strategic layout	Not part of strategic layout	County level strategic layout	City level strategic layout	Provincial strategic layout	National strategic layout
Policy benefits	Implementation of relevant policies within 3 items, with exceedingly significant local limitations in policy influence	3-5 policy benefits, with relatively significant local limitations in policy influence	5-10 policy benefits, with slight local limitations in policy influence	10-15 policy benefits, with slighter local limitations in policy influence	More than 15 favorable policies with significant and wide-ranging impact
High demand that exceeds supply	The supply and demand balance, data supply can meet more than 80% of market demand	The supply and demand are relatively balanced, and the data supply can meet 60% to 80% of the market demand	The supply and demand are relatively unbalanced, and the data supply can meet 40% to 60% of the market demand	The supply and demand imbalance, data supply can meet market demand by 20% to 40%	The supply and demand are extremely unbalanced, and the data supply cannot meet 20% of the market demand
Low demand but precious and scarce	The data resources in the market are abundant and not uncommon, but there is an excess of resources	There are many data suppliers with large data volumes, but not excessive	There are many data providers with large data volumes	There are few data providers and the data volume is relatively small	There are very few data providers, and the data volume is extremely small

(3)Analytic hierarchy process for factor ranking and weighting

Due to the relatively simple structure of the market expectation factor pool and the similar influence of the secondary factors, which can be approximated as equal weighting treatment, this article directly uses the root mean square method of AHP analytic hierarchy process to rank and weight the tertiary factors, and conducts consistency testing.

3.2.5. The Second and Third Level Factor Pools of Cost Value Factor

This article refers to the current industry standard "Evaluation Index System for E-commerce Data Assets" GB/T 37550-2019 and establishes second and third level factor pools for cost value factors. The specific content is shown in the following figure.

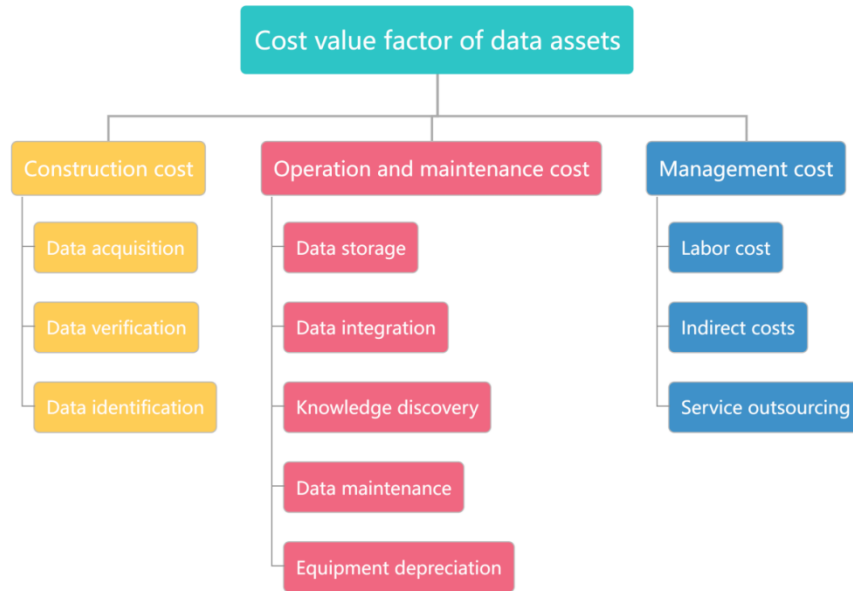


Figure 9: The Second and Third Level Factor Pools of Cost Value Factor.

Enterprises refer to the above factor system to calculate the total cost of their data assets. The calculation formula is as follows:

$$C = \alpha \left(\frac{S}{N} * n \right) \quad (6)$$

S is the total cost of all data, N is the total amount of all data, and n is the effective data volume. To ensure the balance of the influence of the four primary factors, a cost value coefficient $\alpha=0.001\%$ is set.

3.3. Summary

At this point, the multi factor valuation screening and scoring model architecture is completed, divided into first, second, and third level factor pools. It coordinates both subjective and objective dimensions, taking into account objective valuation, audit protection, market expectations, multiple factors of cost value, and multiple parties, meeting the needs of cross industry transactions of data assets while also having good generalization in application scenarios.

At the same time, more comprehensive and accurate valuation and quantification of data assets can be achieved to maximize the benefits for both buyers and sellers. Promote the normative role of third-party audit on data assets, squeeze out the "foam" of data assets, reduce discriminatory analysis, and enhance social effects, while promoting the formation of a benign ecological trading market for data assets.

The overall formula for quantifying comprehensive value:

$$G = \alpha V + S + \gamma E + C \quad (7)$$

V is the quantified value of the objective valuation factor based on effects, $\alpha=10$, S is the quantified value of the security factor combined with auditing, E is the quantified value of the market expectation factor based on AHP analysis, $\gamma=0.1$, and C is the quantified value of the cost value factor.

4. Model Implementation of Water Resources Data Assets in Key Provinces for Comprehensive Management of the Yangtze River

On July 18, 2024, *Decision of the Central Committee of the Communist Party of China on Further Comprehensively Deepening Reform and Promoting Chinese path to modernization* was passed at the 20th Central Committee of the CPC, which points out the need to promote the construction of an integrated ecological environment governance system that connects upstream and downstream in important river basins.

In recent years, the upper and lower reaches of the Yangtze River Basin have broken down barriers and continuously explored cross departmental and cross regional linkage mechanisms. For example, the three provinces of Hubei, Jiangxi, and Hunan jointly signed the "One River, Two Lakes" system governance cooperation agreement, established cross regional coordination mechanisms, joint research mechanisms, and information sharing mechanisms.

4.1. Realistic Demand for horizontal Compensation and Vertical Exchange of Ecological Data Assets in the Yangtze River

In fact, as early as 2021, General Secretary Xi Jinping emphasized at the symposium on further promoting the high-quality development of the Yangtze River Economic Belt that "improving the horizontal ecological protection compensation mechanism and stimulating the enthusiasm of the entire basin to participate in ecological protection".

On the one hand, strengthening horizontal compensation, that is, compensation between watersheds or regions, aims to regulate the interest relationships between entities that do not have a subordinate relationship, explore market-oriented compensation methods, and form a multi-party flow of ecological data assets led by the government and participated by enterprises and all sectors of society; On the other hand, building an integrated ecological environment data sharing and exchange system in the Yangtze River Basin can enhance the data sharing capacity between provinces and cities in the Yangtze River Basin, that is, the data exchange between upstream and downstream governments.

This poses a practical need for the confirmation, valuation, pricing, and even commercialization of data assets. The multi factor valuation screening and scoring model for quantifying the value of data assets in this article can alleviate the current urgent problems, assist in the inventory valuation and value quantification of ecological data assets in the Yangtze River, and promote the circulation and trading of multidimensional entities of data assets.

Currently, in the first year of hydrological data asset inventory, some provinces have started pilot explorations. On August 12, 2024, the Hubei Provincial Hydrological Center issued a notice to implement the provincial government's decision and deployment to deepen the inventory of state-owned "three assets". The center carried out the first "confirmation, ownership confirmation, and value confirmation" work of hydrological data resources in Hubei Province. Establish a specialized team to carry out statistical and valuation work on hydrological data resources throughout the province. However, the inventory of ecological data assets is still in its early stages and has not yet been fully implemented. The multi factor valuation screening model constructed in this article can help promote related work.

4.2. Model Practice

This article selects the five key provinces for comprehensive management of the Yangtze River Basin: Sichuan, Chongqing, Hubei, Anhui, and Jiangsu. Empirical research on quantifying the value of its water resources data assets. The data is sourced from the latest water resources bulletins released by various provincial water conservancy bureaus, namely the "2023 Sichuan Water Resources Bulletin", "2022 Chongqing Water Resources Bulletin", "2022 Hubei Water Resources Bulletin", "2022 Anhui Water Resources Bulletin", and "2023 Jiangsu Water Resources Bulletin".

4.2.1. Objective valuation factor based on utility (V)

Firstly, the 13 secondary factors of water resources data assets in each province were quantified, and the results are shown in Table 6.

Table 6: Quantitative values of secondary factors of objective valuation factor.

Indicator Name	Sichuan Province	Chongqing Province	Hubei Province	Anhui Province	Jiangsu Province
Accuracy (s1)	0.8182	0.8342	0.8756	0.8947	0.8072
Integrity (s2)	0.7664	0.7821	0.8532	0.8324	0.7321
Consistency (s3)	0.8364	0.6143	0.6250	0.5100	0.8287
Timeliness(s4)	0.8854	0.8432	0.8432	0.8432	0.8432
Normalization(s5)	0.9854	0.8651	0.9654	0.9732	0.8556
Efficiency(s6)	0.8032	0.6774	0.7917	0.7895	0.7732
Credibility(s7)	0.9234	0.9234	0.9234	0.9234	0.9234
Precision(s8)	0.8465	0.7963	0.8021	0.8342	0.7987
Recoverability(s9)	0.8654	0.8013	0.8531	0.8246	0.8354
Accessibility(s10)	0.9653	0.9653	0.9653	0.9653	0.9453
Authenticity(s11)	0.9754	0.8571	0.9782	0.9432	0.9342
Data Format (s12)	0.8763	0.8521	0.9124	0.9321	0.8365
Data Performance(s13)	0.8342	0.8683	0.8763	0.8934	0.8921

Subsequently, based on the practical needs of the Yangtze River ecological governance industry, the heterogeneous preferences λ_i of the factors were assigned values, and through the calculation of the utility function, different factors were sorted and weighted accordingly. Make data assets with high final value quantification more in line with the characteristics requirements of the Yangtze River governance industry for data, such as high accuracy and precision for ecological monitoring data.

Table 7: Assignment Results of Heterogeneity Preference λ_i for Yangtze River Ecological Data Factors.

	Accuracy	Integrity	Consistency	Timeliness	Normalization
Utility preference (λ_i)	3	2	∞	1	∞
	Efficiency	Credibility	Precision	Recoverability	Accessibility
Utility preference (λ_i)	2	2	2	-1	∞
	Authenticity	Data Format	Data Performance		
Utility preference (λ_i)	1	-1	2		

Table 8: Results of Utility Function Value Calculation.

Indicator Name	Sichuan Province	Chongqing Province	Hubei Province	Anhui Province	Jiangsu Province
Accuracy (u1)	0.1580	0.1438	0.1072	0.0902	0.1678
Integrity (u2)	0.1910	0.1776	0.1175	0.1347	0.2211
Consistency (u3)	0.1034	0.2739	0.2648	0.3680	0.1087
Timeliness(u4)	0.0707	0.0988	0.0988	0.0988	0.0988
Normalization(u5)	0.0086	0.0840	0.0205	0.0158	0.0904
Efficiency(u6)	0.1594	0.2698	0.1693	0.1711	0.1851
Credibility(u7)	0.0601	0.0601	0.0601	0.0601	0.0601
Precision(u8)	0.1230	0.1654	0.1605	0.1332	0.1632
Recoverability(u9)	0.1992	0.2851	0.2161	0.2545	0.2402
Accessibility(u10)	0.0205	0.0205	0.0205	0.0205	0.0327
Authenticity(u11)	0.0145	0.0894	0.0128	0.0340	0.0396
Data Format (u12)	0.1841	0.2175	0.1327	0.1038	0.2386
Data Performance(u13)	0.1332	0.1050	0.0985	0.0843	0.0856

4.2.2. Security Factor Combined with Audit (S)

Firstly, based on the second and third level security factors, relevant data that does not meet audit requirements will be excluded. Then, the quality of audit security assurance for compliant data was quantified using the Delphi method. Taking Hubei Province as an example, the quantitative results are as follows, and other provinces can refer to this process.

Table 9: The Results of Kendall W Test.

	Kendall harmony coefficient		Kendall W test		
Number of experts	Number of evaluation items		Kendall coordination coefficient	statistic χ^2	p
8	13		0.895	85.920	0.000

Table 10: Quantitative scoring concentration.

term	average value	standard deviation	25% percentile	50% percentile	75% percentile	Coefficient of variation (CV)
T1	3.750	0.463	3.8	4.0	4.0	0.123
T2	3.875	0.641	3.8	4.0	4.0	0.165
T3	4.500	0.535	4.0	4.5	5.0	0.119
T4	7.875	0.354	8.0	8.0	8.0	0.045
T5	7.000	0.535	7.0	7.0	7.0	0.076
T6	6.750	0.463	6.8	7.0	7.0	0.069
T7	5.625	0.518	5.0	6.0	6.0	0.092
T8	7.375	0.518	7.0	7.0	8.0	0.070
T9	4.875	0.641	4.8	5.0	5.0	0.131
T10	5.500	0.535	5.0	5.5	6.0	0.097
T11	5.750	0.707	5.0	6.0	6.0	0.123
T12	5.000	0.535	5.0	5.0	5.0	0.107
T13	4.625	0.518	4.0	5.0	5.0	0.112

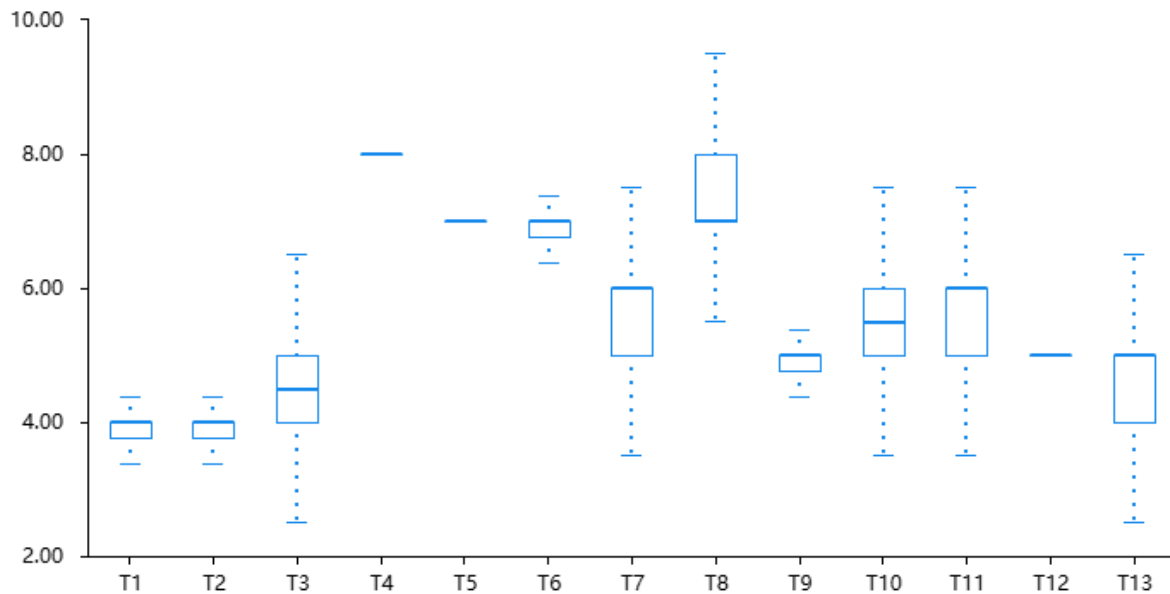


Figure 10: Delphi Method Scoring Boxplot.

Table 11: Summary of Quantitative Results of Security Assurance Factors Combined with Audit in Each Province.

	Sichuan Province	Chongqing Province	Hubei Province	Anhui Province	Jiangsu Province
T1	4.000	3.500	3.750	3.500	4.500
T2	3.333	3.750	3.875	3.750	3.750
T3	3.333	4.750	4.500	4.500	5.500
T4	7.667	7.750	7.875	7.950	7.750
T5	7.000	6.500	7.000	7.250	6.750

Table 11: (continued).

T6	5.333	7.500	6.750	7.750	6.250
T7	5.333	5.750	5.625	5.500	5.250
T8	7.333	7.250	7.375	7.250	7.250
T9	4.333	4.500	4.875	4.500	4.500
T10	4.667	5.250	5.500	5.250	5.250
T11	5.333	4.750	5.750	5.250	4.750
T12	3.667	5.500	5.000	5.750	3.750
T13	4.333	4.500	4.625	4.750	4.750

After completing the factor quantification of the relevant tertiary factors, the factors were sorted and weighted again using the Compound AHP Analytic Hierarchy Process by SPSSAU software, and the combined weights of the tertiary factors were obtained.

Step one, assign weights and consistency checks to the target layer (secondary factors) using the root mean square AHP analysis method:

Table 12: The judgment matrix A of the target layer.

	B1	B2	B3	B4	ω_i
B1	1.000	0.200	0.143	0.333	5.143%
B2	5.000	1.000	0.200	3.000	21.667%
B3	7.000	5.000	1.000	6.000	62.671%
B4	3.000	0.333	0.167	1.000	10.519%

Table 13: Summary of Consistency Inspection Results.

Maximum eigenvalue	CI	RI	CR	Consistency test result
4.236	0.079	0.890	0.088	Pass

The calculated CI value for the 4th order judgment matrix here is 0.079, and the RI value is 0.890 according to the table. Therefore, the calculated $CR=0.088<0.1$, which means that the judgment matrix in this study meets the consistency test and the calculated weights have consistency.

Step two, perform root mean square AHP analysis weighting and consistency testing on the indicator layer (C_i) of each target layer (B_i):

Table 14: Judgment Matrix B1 and Weighting Results.

B1	C1	C2	C3	ω_i
C1	1.000	0.500	0.333	16.342%
C2	2.000	1.000	0.500	29.696%
C3	3.000	2.000	1.000	53.961%

B1 judgment matrix $\lambda_{\max}=3.009, CI=0.005, RI=0.520, CR=0.009<0.1$, meet the consistency check.

Table 15: Judgment Matrix B2 and Weighting Results.

B2	C4	C5	ω_i
C4	1.000	2.000	66.667%
C5	0.500	1.000	33.333%

B2 judgment matrix $\lambda_{\max} = 2.000$, $CI = 0.000$, $RI = 0.000$ (unable to calculate CR value due to $RI = 0$), but the second-order data all meet the consistency test.

Table 16: Judgment Matrix B3 and Weighting Results.

B3	C6	C7	C8	C9	C10	ω_i
C6	1.000	5.000	0.500	6.000	7.000	38.982%
C7	0.200	1.000	0.500	2.000	3.000	13.876%
C8	2.000	2.000	1.000	2.000	5.000	32.139%
C9	0.167	0.500	0.500	1.000	3.000	10.139%
C10	0.143	0.333	0.200	0.333	1.000	4.864%

B3 judgment matrix $\lambda_{\max} = 5.399$, $CI = 0.100$, $RI = 1.120$, $CR = 0.089 < 0.1$, meet the consistency check.

Table 17: Judgment Matrix B4 and Weighting Results.

B4	C11	C12	C13	ω_i
C11	1.000	0.500	2.000	31.081%
C12	2.000	1.000	2.000	49.339%
C13	0.500	0.500	1.000	19.580%

B4 judgment matrix $\lambda_{\max} = 3.054$, $CI = 0.027$, $RI = 0.520$, $CR = 0.052 < 0.1$, meet the consistency check.

Step three, combine and weight the results of the target layer (B_i) and indicator layer (C_i) mentioned above, and the results are as follows:

Table 18: The Combined Weighting Results of The Target Layer (B_i) and Indicator Layer (C_i).

Administrative Levels	B1 (0.05143)	B2 (0.21667)	B3 (0.62671)	B4 (0.10519)	Combined Weight (ω_i)
C1	0.16342	0	0	0	0.0084
C2	0.29696	0	0	0	0.0153
C3	0.53961	0	0	0	0.0278
C4	0	0.66667	0	0	0.1444
C5	0	0.33333	0	0	0.0722
C6	0	0	0.38982	0	0.2443
C7	0	0	0.13876	0	0.0870
C8	0	0	0.32139	0	0.2014
C9	0	0	0.10139	0	0.0635
C10	0	0	0.04864	0	0.0305
C11	0	0	0	0.31081	0.0327
C12	0	0	0	0.49339	0.0519
C13	0	0	0	0.1958	0.0206

4.2.3. Market Expectation Factor Based on AHP Analysis (E)

Based on the detailed rules of the subjective scoring model for data assets, this article subjectively scored various factors related to market expectations for the Yangtze River ecological data assets. The specific results are as follows:

Table 19: Third Level Factor Scoring Results of Market Expectation Factor Pool for Yangtze River Ecological Data Assets.

Factor name	Profit preference (C1)	Risk preference (C2)	Cash flow inflow (C3)	Strategic layout (C4)
Score	15	10	78	90
Factor name	Policy benefits (C5)	High demand that exceeds supply (C6)	Low demand but precious and scarce (C7)	
Score	92	75	55	

After completing the factor quantification of the relevant tertiary factors, the root mean square method of AHP was used to rank and weight the tertiary factors. The results of the weighting and consistency test are as follows:

Table 20: The AHP Judgment Matrix And Weighting Results of The Three-level Factors of Market Expectation Factor.

	C1	C2	C3	C4	C5	C6	C7	ω_i
C1	1.000	0.333	3.000	0.143	0.200	0.143	0.143	3.324%
C2	3.000	1.000	3.000	0.200	0.200	0.200	0.333	5.653%
C3	0.333	0.333	1.000	0.143	0.143	0.143	0.250	2.507%
C4	7.000	5.000	7.000	1.000	3.000	3.000	5.000	36.406%
C5	5.000	5.000	7.000	0.333	1.000	0.500	3.000	18.244%
C6	7.000	5.000	7.000	0.333	2.000	1.000	3.000	23.335%
C7	7.000	3.000	4.000	0.200	0.333	0.333	1.000	10.530%

Table 21: Summary of Consistency Inspection Results.

Maximum eigenvalue	CI	RI	CR	Consistency test result
7.674	0.112	1.360	0.083	Pass

The calculated CI value for the 4th order judgment matrix is 0.059, and the RI value is 0.890 according to the table. Therefore, the calculated CR value is $0.066 < 0.1$, which means that the judgment matrix in this study meets the consistency test and the calculated weights have consistency.

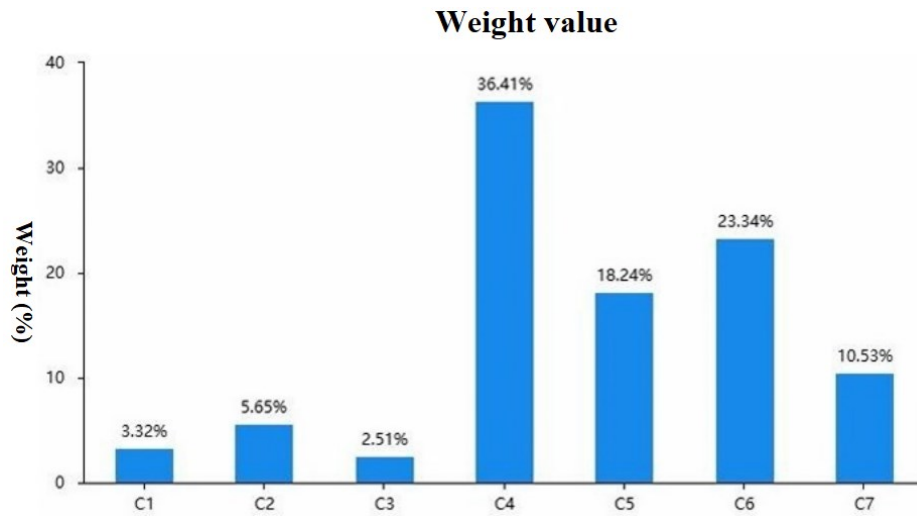


Figure 11: Bar Chart of Factor Weights.

4.2.4. Cost Value Factor (C)

Due to the fact that data asset inventory and valuation are still in their early stages, the cost and investment of constructing, operating, and managing water resource data warehouses in various provinces have not been accurately disclosed by the water resources departments. This article quantifies the cost value factors of water resource data assets in each province based on the annual average cost of database construction, operation, and management in the market as of August 2024 and the effective water resource data volume in each province. When quantifying the value of data assets in various industries and enterprises using this model, it can be calculated based on its own accurate cost statistics.

Table 22: The Quantitative Results of Cost Value Factor of Water Resources Data Assets in Each Province.

	Sichuan Province	Chongqing Province	Hubei Province	Anhui Province	Jiangsu Province
Quantitative value of cost value	2.9175	7.4263	3.8193	8.5668	3.4214

4.3. Summary and Conclusion

The comprehensive quantitative value G of each province was calculated using the overall formula of comprehensive value quantification, and the results are shown in the table below.

Table 23: Quantitative Calculation Results of Comprehensive Value in Each Province.

	Sichuan Province	Chongqing Province	Hubei Province	Anhui Province	Jiangsu Province
Quantitative value of Comprehensive Value (G)	24.7683	29.4465	26.2039	31.1069	25.3351
Rank	5	2	3	1	4

According to the table above, the value of water resources data assets in Anhui Province is the highest, mainly due to its high accuracy, completeness, precision, and standardization of data assets, as well as higher audit quality than other provinces. At the same time, due to the large amount of effective data, the operation and maintenance costs are relatively high.

Compared with other provinces, the value of water resources data assets in Sichuan Province is relatively low. Although its data assets perform well in timeliness, efficiency, and accuracy, their completeness, accuracy, and data performance are poor. At the same time, due to the existence and completeness of data assets, as well as poor data governance structure and organizational performance, its audit quality is not high. In addition, its effective data volume is small, and the cost value of assets is relatively low.

The comprehensive value quantification of this model combines multidimensional factors and multiple stakeholders, taking into account the industry data preferences and requirements of the Yangtze River ecological governance industry. It comprehensively values and quantifies the value of data assets, providing a prerequisite guarantee for the subsequent multi-party trading and circulation of ecological data assets.

5. Conclusion

Effective management and utilization of data assets to maximize revenue cannot be achieved without effective data governance. The end of data governance is the beginning of data management. Starting from January 1, 2024, the "Interim Provisions on Accounting Treatment of Enterprise Data Resources" issued by the Chinese Ministry of Finance officially came into effect, initiating the "first year of data asset entry into the balance sheet". Data assets are embarking on a new journey of resource utilization, commercial circulation, and financial revenue generation. However, due to the fact that the valuation, inclusion, and pricing of data assets are still in their early stages, problems such as discriminatory analysis of data assets, lack of a generalized objective valuation system and a single objective quantitative dimension, are constantly emerging.

At the same time, with the explosive growth of data and industry integration, cross industry data asset circulation is becoming increasingly frequent, and there is an urgent need for a scenario generalization valuation model that can meet the heterogeneous preferences and requirements of different industries and buyers and sellers.

This article seizes the opportunity of the times and, guided by practical problems, constructs a multi factor valuation screening and scoring model that takes into account both subjective and objective dimensions, coordinates multiple market entities such as buyers, sellers, and third-party auditors, and combines multiple factors. To comprehensively and accurately value and quantify the value of data assets, laying the foundation for the utilization and circulation of data assets, as well as the construction of data asset trading markets. By combining and weighting factors through various methods, the quantification of data asset value can be effectively applied in specific scenarios, while also meeting the needs of data asset trading across scenarios and industries.

This model provides new ideas and academic suggestions for the comprehensive valuation and value quantification of data assets, and serves as a reference for the valuation and pricing of data assets in multiple dimensions and scenarios, making certain contributions to the marketization of data assets

References

- [1] Wang Qi.. *Breaking Down the Barriers of Data Exchanges and Building a Unified Data Trading Market -- Based on Comparative Analysis of Aata Exchanges[J]*. *China Telecommunications Trade*,2023,(01):30-33.

- [2] Jiang Qi., Sun Chaochen., Ni Shuo.. *Data Circulation and Value Creation in the Process of Marketization of Data Elements -- A Case Study Based on the Shanghai Data Exchange*[J]. *Research of Finance and Education*,2024,37(03):3-10.
- [3] Even A., Shankaranarayanan G.. *Utility-driven assessment of data quality*[J]. *ACM SIGMIS Database*,2007(2):75 ~ 93.
- [4] Li Y., Chao X., Ercisli S.. *Disturbed-entropy:A simple data quality assessment approach*[J]. *ICF Express*,2022(3):309 ~ 312.
- [5] Yang J., Lan G., Li Y., et al.. *Data quality assessment and analysis for pest identification in smart agriculture*[J]. *Computers and Electrical Engineering*, 2022(103):108322.
- [6] Spiekermann M.. *Data marketplaces:Trends and monetisation of data goods*[J]. *Intereconomics*, 2019(4):208 ~ 216.
- [7] Kong Min., Wang Fengfan., Geng Shanshan., et al.. *Application of Fuzzy Analytic Hierarchy Process in Quality Evaluation of Sub-bottom Profile Data*[J]. *Marine Sciences*,2023,47(10):65-75.
- [8] Sangseop Lim., Chang Hee Lee., Jae Hwan Bae., Young Hun Jeon.. *Identifying the Optimal Valuation Model for Maritime Data Assets with the Analytic Hierarchy Process (AHP)*[J]. *Sustainability. Volume 16 , Issue 8 . 2024*
- [9] Tang Yong., Li Dongpeng., Lin Juanjuan.. *Exploration of Comprehensive Evaluation Method for Data Quality Based on Utility*[J]. *Finance and Accounting Monthly*,2024,45(16):110-116.
- [10] Yizhuo Cheng., Mei Zhou.. *Selection of Methods for Valuing Data Assets -- Taking the Telecommunications Industry as an Example*[J]. *Scientific Journal of Economics and Management Research. Volume 6 , Issue 3 . 2024*
- [11] Song Mufan.. *Research on the Quality Evaluation of Government Open Data from the Perspective of Data Openness*[J]. *Journal of Library and Information Science*,2022,7(08):63-70.
- [12] Wang Jingyong., Wang Shengdan., Xue Lida.. *The Logical Framework and Implementation Mechanism of Enterprise Data Asset Audit*[J]. *Finance and Accounting Monthly*,2024,45(16):90-96.
- [13] Zheng Zhihang.. *Ethical Crisis and Legal Regulation of Artificial Intelligence Algorithms*[J]. *Science of Law - Northwest University of Political & Law*,2021(1):14~26.
- [14] Nespoli A., Ogliari E., Pretto S.,et al.. *Data quality analysis in day-ahead load forecast by means of LSTM*[C]. *2020 IEEE International Conference on Environment and Electrical Engineering and 2020 IEEE Industrial and Commercial Power Systems Europe(IEEEIC/I&CPS Europe)*,2020.
- [15] Tang Chen., Yu Qiancheng., Li Xiaoning.,Lu Zekun., Yang Yufan .. *Application of DeepForest-CQP multi-factor model in quantitative stock selection*[J]. *Journal of Intelligent & Fuzzy Systems. Volume 44 , Issue 3 . 2023. PP 5425-5436*