Research on Maturity Evaluation of Energy Internet Big Data Ecosystem Based on Mathematical Statistical Model

Haixu Song(✉), Xiaofeng Zhan, and Rui Li

Department of Corporate Strategy, State Grid Energy Research Institute Co., Ltd., Beijing, China
shxstudent@126.com, {zhangxiaofeng,lirui}@sgeri.sgcc.com.cn

Abstract. The energy revolution and the digital revolution are integrated deeply. The creation of an Energy Internet big data ecosystem has gradually become a trend. The paper builds an evaluation system for the maturity of the Energy Internet big data ecosystem from the dimensions of ecological subject, key data, and overall environment. And then it uses big data capture, grade evaluation, tomographic analysis, mathematical statistics and other models, to propose the evaluation method of the Energy Internet big data ecosystem maturity. Finally, it takes Region A as an example to carry out an empirical analysis on the maturity of its energy internet big data ecosystem. The analysis found that the maturity of the energy Internet big data ecosystem in Region A is generally average and needs to be improved. Especially in terms of key data, there are still major deficiencies in data types and data quality. The degree of data sharing and mining has great room for improvement. The paper can provide reference for the digital transformation of energy and power companies and the energy industry innovative development.

Keywords: Energy Internet · big data · ecosystem · maturity · mathematical statistics model · evaluation system

1 Introduction

In recent years, emerging digital technologies such as big data, cloud computing, and the Internet of Things have developed vigorously, bringing about all-round changes in thinking, technology, and models. Data has become a new factor of production and is changing the way of economic development. It has become the key to the development of new drivers of traditional industries to form data assets through digital value mining. With the acceleration of global energy reforms, large-scale development of new energy, and rapid progress in new technologies, cloud computing and big data technologies are gradually being integrated with energy technologies [1]. As the key energy companies that are related to the national economy and people's livelihood, power grid companies are transforming and upgrading to the Energy Internet. Big data technology can also be conducive to the realization of intelligent management of power enterprises [4]. In recent years, the Energy Internet and big data have been integrated deeply. The creation of an Energy Internet big data ecosystem has gradually become a hot spot in the academic

© The Author(s) 2023
Z. Zhan et al. (Eds.): ICBDIE 2022, AHCS 5, pp. 294–304, 2023.
https://doi.org/10.2991/978-94-6463-034-3_30

and business circles. How to evaluate the maturity of the ecosystem, and then find the shortcomings to improve it, has become an important issue that needs to be studied in the new situation.

However, in the existing technology, there is little relevant analysis method for judging and evaluating the Energy Internet big data ecosystem maturity in a certain region. Some scholars have conducted in-depth research on the ecology of the energy industry, but few have combined research with big data. Some scholars have conducted in-depth research on big data, but they rarely combine the characteristics of the energy industry. Therefore, this paper integrates the energy industry and data ecology deeply, and conducts in-depth research on the Energy Internet big data ecosystem. It can provide reference for the digital transformation of energy and power companies and the energy industry innovative development.

2 Basic Framework of the Energy Internet Big Data Ecosystem

2.1 Basic Theory

Energy Internet is the product of the deep integration of energy system and internet technology. It is a new generation of energy system with electricity as the hub and platform. Energy Internet is one of the key technologies of the new round of energy reform. Its development goal is to build a green, low-carbon, safe, efficient, and open and shared energy ecosystem [5].

Enterprise ecosystem theory is an open and systematic theory based on management, ecology and system theory. The enterprise ecosystem draws resources from the environment, passes it on and transforms it, and finally realizes value-added through processing and allocating them [2].

Big data is a data collection characterized by large capacity, multiple types, fast access speed, and high application value. Big data has become an important strategic resource, which is of great significance to the development of existing industries and science and technology [3].

2.2 Basic Framework

This paper construct the Energy Internet big data ecosystem model mainly based on the theory of ecosystems, combining with the energy industry chain and data industry chain. Finally, the ecosystem model includes three levels: core subjects, relevant subjects, and macro-environment.

1) Core subjects
   The core subject layer is built around the core industrial chain of the Energy Internet. The Energy Internet mentioned in this paper refers to the integration of advanced information and communication technology, control technology and advanced energy technology with electricity as the center and a strong smart grid as the basic platform. The Energy Internet can support clean and low-carbon transformation and comprehensive utilization of energy. Multiple subjects can access

to the system efficiency optimization and flexibly and conveniently. According to energy production, transmission, consumption, storage, consumption, management and other links, the core subjects mainly include grid companies, energy suppliers, energy consumers, energy storage suppliers, equipment suppliers, and government departments.

2) Relevant subjects
The relevant subject layer is built around the related derivative industries of the Energy Internet. In addition to the core industrial chain, the Energy Internet also requires financial services, technical services, data services, information services, asset services and other services. So various ecological entities are derived, including technical service providers, financial service providers, data service providers, communication operators, operation and maintenance service providers, traditional competitors, cross-border competitors, industry associations and think tanks, etc.
3) Macro-environment
The environment layer is the macro environment on which the main subjects of the Energy Internet industry chain rely for survival. According to the PEST model, combined with the characteristics of the Energy Internet business, the environmental layer is divided into political environment, economic environment, industrial environment, technological environment, social environment, and international environment.

3 Study on the Evaluation System of Energy Internet Big Data Ecosystem Maturity

3.1 Evaluation Index System

The evaluation system of the Energy Internet big data ecosystem maturity includes three dimensions: ecological subjects, key data, and overall environment.

Ecological subjects include three aspects: subject diversity, degree of correlation between subjects, and subjects' willingness to participate in the ecosystem. Subject diversity refers to whether it exists various related subjects of the Energy Internet in the ecosystem. The degree of correlation between subjects refers to whether the core subjects and the relevant subjects are closely related. The subjects' willingness to participate in the ecosystem refers to whether the core and relevant subjects are willing to join the Energy Internet big data ecosystem, or whether the willingness is strong.

Key data includes data diversity, data objectivity, data sharing degree, and data mining degree. Data diversity refers to whether the data in the ecosystem covers the energy industry chain and multiple types of data such as technology, policy, environment, finance, customers, etc. Data objectivity refers to the authenticity, validity and timeliness of the data. The degree of data sharing refers to whether the various data can be multi-directionally flowed and fully shared among the core subjects and related subjects. The degree of data mining refers to whether the ecosystem has complete software, hardware and technology for big data analysis and processing, to realize the full mining of data value and form a mature business model.

The overall environment includes economic level, policy intensity, and innovation atmosphere. The economic level refers to whether the economic development level in

Table 1. The Evaluation Index System of Energy Internet Big Data Ecosystem Maturity.

Dimension	Sub-dimensions	Main evaluation content
Ecological subjects	Subject diversity	Whether it exists various related subjects of the Energy Internet in the ecosystem
	Degree of correlation between subjects	Whether the core subjects and the relevant subjects are closely related
	Subjects' willingness to participate in the ecosystem	Whether the core and relevant subjects are willing to join the Energy Internet big data ecosystem, or whether the willingness is strong
Key data	Data diversity	Whether the data in the ecosystem covers the energy industry chain and multiple types of data such as technology, policy, environment, finance, customers, etc.
	Data objectivity	The authenticity, validity and timeliness of the data
	Data sharing degree	Whether the various data can be multi-directionally flowed and fully shared among the core subjects and related subjects
	Data mining degree	Whether the ecosystem has complete software, hardware and technology for big data analysis and processing, to realize the full mining of data value and form a mature business model
Overall environment	Economic level	Whether the economic development level in the region can support the construction of an energy big data ecosystem, and whether the infrastructure such as communications is complete
	Policy intensity	Whether the region has complete energy and big data policies which can support the healthy development of the energy big data ecology
	Innovation atmosphere	Whether the region has an active big data innovation atmosphere that can create a good environment for the construction of the energy big data ecosystem

the region can support the construction of an energy big data ecosystem, and whether the infrastructure such as communications is complete. The intensity of the policy refers to whether the region has complete energy and big data policies which can support the healthy development of the energy big data ecology. The innovation atmosphere refers to whether the region has an active big data innovation atmosphere that can create a good environment for the construction of the energy big data ecosystem (Table 1).

3.2 System Evaluation Method

Through the Energy Internet big data ecosystem maturity evaluation index system, we can find that many evaluation content is qualitative content, which is difficult to be quantified directly. So the paper evaluates the ecosystem maturity using questionnaire survey method, AHP, and expert evaluation method, which specifically includes five steps: designing questionnaires, determining levels and standards, data labelling, setting weights, and analyzing results.

1) Designing questionnaires
   According to the Energy internet big data ecosystem maturity evaluation index system, it designs several questions for each sub-dimension.
2) Determining levels and standards
   Through the likert table, it determine the standard of each grade, which is divided into five grades by 1–5 points respectively: poor, average, good, very good, and perfect.
3) Data labelling
   According to the different attributes of each questionnaire, each questionnaire is labelled to facilitate subsequent analysis from the two aspects of the region and the time. First, according to the regional location of the ecosystem, it can be divided into three regions: the eastern region, the central region, and the western region. Secondly, according to the formation time, the questionnaires are marked for trend analysis. Assuming a total of 200 valid questionnaires, the labelling results are shown as follows.
4) Setting weight
   It uses the analytic hierarchy process to set the weight of each dimension. Firstly, it builds a hierarchy model. The of the Energy Internet big data ecosystem maturity is the target layer, the dimensions of ecological subjects, key data, and overall environment are the middle layer, and the sub-dimensions such as the diversity of subjects are the factor layers.

Secondly, it constructs the consistent matrix method. All factors are compared with each other in order to minimize the difficulty of comparing factors with different properties. The scaling method of the judgment matrix a_{ij} is shown as follows.

Thirdly, a consistency check is performed. Consistency index $CI = (\lambda\text{-}n)/(n\text{-}1)$, where λ is the largest eigenvalue of the matrix. In order to measure the size of CI, a random consistency index RI is introduced, and the RI table can be queried. Consistency ratio $CR = CI/RI$.

Layer A has m factors $A_1, A_2, \ldots, A_m$, and the order of the total target Z is $a_1, a_2, \ldots, a_m$.

Layer B has n factors $b_{1j}, b_{2j}, \ldots, b_{nj}$ ($j = 1, 2, 3, \ldots, m$), which are ranked for the hierarchical single order of the factor A_j in the upper layer A.

The total hierarchical order of layer B (the weight of the i factor of layer B to the total target is $\sum_{j=1}^{m}(a_j b_{ij})$.

B_1: $a_1 b_{11} + a2b_{12} + \ldots + a_m b_{1m}$

B_2: $a_1 b_{21} + a2b_{22} + \ldots + a_m b_{2m}$

Table 2. Questionnaire Survey of Energy Internet Big Data Ecosystem Maturity Evaluation.

Questions
Ecological subjects
Subject diversity
Q1: Whether it exists core subject such as grid companies, energy suppliers, energy consumers, energy storage suppliers, and government departments, etc.?
Q2: Whether it exists relevant subject such as technical service providers, financial service providers, data service providers, communication operators, operation and maintenance service providers, traditional competitors, cross-border competitors, industry associations and think tanks, etc.?
Degree of correlation between subjects
Q3: How is the degree of connection among the core subjects?
Q4: How is the degree of connection among the relevant subjects?
Q5: How is the degree of connection among the core subjects and relevant subjects?
Subjects' willingness to participate in the ecosystem
Q6: How is the willingness of core subjects to participate in the ecosystem?
Q7: How is the willingness of relevant subjects to participate in the ecosystem?
Key data
Data diversity
Q8: Can data cover all aspects of energy production, transmission, consumption, storage, and management?
Q9: Does the data include multiple types of data such as production, technology, policy, environment, finance, and customers?
Data objectivity
Q10: How is the authenticity of the data?
Q11: How is the validity of the data?
Q12: How is the timeliness of the data?
Data sharing degree
Q16: Can the data of the core subjects be fully shared among core subjects?
Q17: Can the data of the core subjects be fully shared among relevant subjects?
Q18: Can the data of the relevant subjects be fully shared among relevant subjects?
Q19: Can the data of the relevant subjects be fully shared among core subjects?
Data mining degree
Q20: Does it has data mining related facilities and technologies such as artificial intelligence and cloud computing?
Q21: Does it have mined data value fully and formed a mature business model?

(*continued*)

Table 2. (*continued*)

Questions
Overall environment
Economic level
Q22: How is the local economic level?
Q23: Is the communication infrastructure such as the internet and 5G complete?
Policy intensity
Q24: Does the local government support the energy industry and big data development?
Q25: How complete is the region's laws and regulations regarding data security, data ownership?
Innovation atmosphere
Q26: What is the overall level of local technological innovation?
Q27: Do the overall innovation environment and atmosphere of the region fully encourage innovation?

Table 3. Determining levels and standards of Energy Internet Big Data Ecosystem Maturity.

Score	Evaluation basis		
	Ecological subjects	Key data	Overall environment
1	Ecological subject is relatively single. The degree of correlation between subjects is weak. The willingness of subjects to participate is weak.	The data type is relatively single. The data objectivity is weak. The degree of sharing and mining is very insufficient.	The local economy is underdeveloped. The policy support is very weak. The innovation atmosphere is very weak.
2	Ecological subjects is a little single. The degree of correlation between subjects is weak. The willingness of subjects to participate is not strong.	The data type is a little single. The objectivity is not strong. The degree of sharing and mining is insufficient.	The local economy is underdeveloped. The policy support is weak. The innovation atmosphere is weak.
3	There are many kinds of ecological subjects. The degree of correlation between subjects is acceptable. The willingness of subjects to participate is acceptable.	There are various types of data. The objectivity is acceptable. The degree of sharing and mining is acceptable.	The local economic level is good. The policy support is acceptable. The innovation atmosphere is acceptable.

(*continued*)

Table 3. (*continued*)

Score	Evaluation basis		
	Ecological subjects	Key data	Overall environment
4	Ecological subjects are more abundant. The degree of correlation between subjects is relatively close. The subjects are willing to participate.	There are various types of data. The objectivity is acceptable. The degree of sharing and mining is acceptable.	The local economic level is good. The policy support is strong. The innovation atmosphere is good.
5	There are many ecological subjects. The degree of correlation between subjects is very close. The willingness of subjects to participate is very strong.	The data types are very rich. The objectivity is very strong. The degree of sharing and mining is very strong.	The local economic level is very good. The policy support is very strong. The innovation atmosphere is very good.

Table 4. Questionnaire information labelling processing.

Region	Eastern		Central		Western	
Label	R1		R2		R3	
Number	85		65		50	
Time	2020.1	2020.6	2021.1	2021.6	2022.1	…
Label	T01	T02	T11	T12	T21	…
Number	40	40	40	40	40	…

…

B_n: $a_1 b_{n1} + a2 b_{n2} + \ldots + a_m b_{nm}$

The consistency ratio of the hierarchical total ranking is:

$CR = (a_1 CI_1 + a_2 CI_2 + \ldots + a_m CI_m)/(a_1 RI_1 + a_2 RI_2 + \ldots + amRI_m)$.

When $CR < 0.1$, the hierarchical total ranking has passed the consistency test. Otherwise, the matrix A should be reconstructed and a_{ij} should be adjusted.

5) Analyzing data

It selects a certain number of experts who have in-depth understanding of energy development, Energy Internet construction, energy big data, data mining, etc., to mark the questions. According to the score of each question, it can calculate the score of each sub-dimension. According to the weight of each dimension, the overall score is obtained (Tables 2, 3, 4, 5, 6 and Fig. 1).

Table 5. The scaling method of the judgment matrix.

Scaling	Meaning
1	Two factors are of equal importance.
3	One factor is slightly more important than the other.
5	One factor is significantly more important than the other.
7	One factor is strongly more important than the other.
9	One factor is extremely important than the other.
2,4,6,8	The median of the above two adjacent judgments.

Table 6. Energy internet big data ecosystem Maturity in Region A

Evaluation dimension		Score for each dimension	Over score
Ecological subjects (0.35)	Subject diversity (0.35)	3.6	3.27
	Degree of correlation between subjects(0.4)	3.4	
	Subjects' willingness to participate in the ecosystem(0.25)	2.6	
Key data (0.4)	Data diversity (0.25)	2.8	2.02
	Data objectivity (0.2)	2.3	
	Data sharing degree (0.3)	1.6	
	Data mining degree (0.25)	1.5	
Overall environment (0.25)	Economic level (0.4)	3.5	3.17
	Policy intensity (0.3)	3.2	
	Innovation atmosphere (0.3)	2.7	

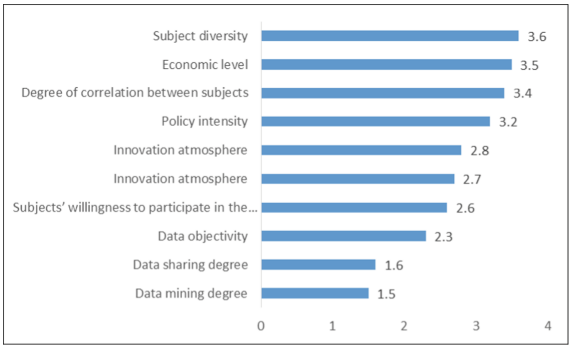

Fig. 1. Energy internet big data ecosystem Maturity in Region A.

Overall score = Σ score of each dimension *the weight of each dimension.

Finally, according to the overall score and each dimension score, we can evaluate the maturity of the Energy Internet big data ecosystem, and analyze the main advantages and disadvantages. At the same time, regional analysis and practice change analysis can also be carried out.

4 Empirical Analysis

It selects the Energy Internet big data ecosystem in Region A for evaluation. According to the evaluation index system, we choose 15 experts in the fields of Energy Internet, big data, power system and other fields from government departments, enterprises, universities, scientific research units and other institutions to evaluate the Energy Internet big data ecosystem maturity. Finally, the weights of ecological subjects, key data and the overall environment are divided into 0.35, 0.4 and 0.25. The comprehensive score of the Energy Internet big data ecosystem maturity in Region A is 2.74, which needs to be strengthened.

In terms of sub-dimensions, the maturity of the ecological subject is the highest, reaching 3.27. This is mainly due to the rapid development of the energy industry. At present, the energy industry has established a complete chain of energy production, transmission, consumption, storage, etc. The industrial chain is complete, and the ecological subjects are diverse and closely linked. Critical data has the lowest maturity at 2.02. It shows that although the energy industry has established a basically complete ecological relationship, there are still major deficiencies in terms of data, especially the degree of data sharing and mining. And this is also the key focus of improving the Energy Internet big data ecosystem in the future. The maturity score of the overall environment is 3.17, not very high. It mainly due to the relatively good level of economic development and strong policy support in Region A, which can provide favourable conditions for the development of innovative models. However, the innovation atmosphere is not very good. The initiative and enthusiasm of various entities still need to be improved, and the substantive business still needs to be explored and developed.

5 Conclusions

The energy revolution and the digital revolution are integrated deeply. The creation of an Energy Internet big data ecosystem has gradually become a trend. The paper builds an evaluation system for the maturity of the Energy Internet big data ecosystem from the dimensions of ecological subject, key data, and overall environment. And then it uses big data capture, grade evaluation, tomographic analysis, mathematical statistics and other models, to propose the evaluation method of the Energy Internet big data ecosystem maturity. Finally, it takes Region A as an example to carry out an empirical analysis on the maturity of its energy internet big data ecosystem. The analysis found that the maturity of the energy Internet big data ecosystem in Region A is generally average and needs to be improved. Especially in terms of key data, there are still major deficiencies in data types and data quality. The degree of data sharing and mining has great room for improvement. The paper can provide reference for the digital transformation of energy and power companies and the energy industry innovative development (Fig. 1).

Acknowledgements. This work is supported by Science and Technology Foundation of SGCC (5400-202018223A-0-0-00).

References

1. TIAN, S. M., LUAN, W. P., ZHANG D. X., etc. (2015). Energy internet technology form and key technologies. Proceedings of the CSEE. 35(14): 3482-3494.
2. WAN, C. (2015). Research on the model and operation mechanism of enterprise ecosystem. China Business and Trade. 7: 183-185.
3. WANG, J. Y. (2020). Analysis of China's big data industry situation in 2020. Software and Integrated Circuits. 10: 76-87
4. YIN, W. L. (2018). The development trend of electric power big data under the situation of energy internet. Computer Products and Circulation. 10, 64.
5. ZHOU, X. X., ZENG, R., GAO, F., QU, L. (2017). Development status and prospect of energy internet. SCIENTIA SINICA Informationis. 47(2): 149-170.

Open Access This chapter is licensed under the terms of the Creative Commons Attribution-NonCommercial 4.0 International License (http://creativecommons.org/licenses/by-nc/4.0/), which permits any noncommercial use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons license and indicate if changes were made.

The images or other third party material in this chapter are included in the chapter's Creative Commons license, unless indicated otherwise in a credit line to the material. If material is not included in the chapter's Creative Commons license and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder.